Watch and learn

# monitoring and reflection to improve the quality of physical therapy care

Simone van Dulmen

Watch and learn

# Monitoring and reflection to improve the quality of physical therapy care

Simone van Dulmen

For reasons of consistency within this thesis, some terms have been standardized throughout the text. As a consequence the text may differ in this respect from the articles that have been published.

The work presented in this thesis was carried out within the Radboud Institute for Health Sciences, at the department Scientific Center for Quality of Healthcare (IQ healthcare), Radboud university medical center.

Publication of this thesis was financially supported by the Scientific College Physical Therapy (WCF) of the Royal Dutch Society for Physical Therapy (KNGF) and the Radboud University.

The studies in this thesis were financially supported by the Ministry of Health, Welfare and Sport (VWS), the Royal Dutch Society of Physical Therapy (KNGF), the Dutch Council for Quality of Healthcare (Regieraad Kwaliteit van Zorg), and health insurance company CZ.

**ISBN**
978-90-9029707-1

**Cover art**
BYRON, Nijmegen

**Design/lay-out**
Promotie In Zicht, Arnhem

**Print**
Ipskamp Printing, Enschede

Watch and learn

# Monitoring and reflection to improve the quality of physical therapy care

**Proefschrift**

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 8 juni 2016
om 14.30 uur precies

door

**Simone Antoinet van Dulmen**
geboren op 7 februari 1975
te Deventer

### Promotor
Prof. dr. M.W.G. Nijhuis-van der Sanden

### Copromotoren
Dr. J.C.C. Braspenning
Dr. J.B. Staal
Dr. P.J. van der Wees

### Manuscriptcommissie
Prof. dr.  W.J.J. Assendelft
Prof. dr. R.A. de Bie (Universiteit Maastricht)
Prof. dr. C. Veenhof (UMC Utrecht)

# Contents

# 1

# General introduction

This thesis describes the process of identifying factors that influence the quality of physical therapy in primary care. This chapter presents essential background information for the different studies involved in that process. It addresses the central elements in care that can be improved. The chapter concludes with an outline of the thesis.

## Quality of Care

Quality of care has been defined by the Institute of Medicine (IOM) in the United States as *"doing the right thing, at the right time, in the right way, for the right person, and having the best possible results*."[1] To enhance the quality of care, the IOM established a framework with recommended improvements for achieving high quality care in the following six areas of healthcare:

- Safety: Efforts in this area should focus on ensuring that patients are not harmed by the care that is intended to help them.
- Effectiveness: This entails providing interventions based on the latest scientific knowledge to all patients who may benefit. As medical knowledge advances, so should the care delivered to patients.
- Patient centeredness: These efforts call for care provision that is responsive to individual patient preferences, needs, and values, and that ensures that patient values guide all clinical decisions. The patient is an integral part of the care team and should be involved in the collaboration e between care providers in making clinical decisions.
- Timeliness: Efforts here aim at reducing waits and sometimes harmful delays for both those who receive and those who give care.
- Efficiency: This involves avoiding waste, including waste of equipment, supplies, ideas, and energy.
- Equity: These efforts focus on providing care that does not vary in quality because of personal characteristics, such as gender, ethnicity, geographic location, and socio-economic status.

This manuscript will focus on the three dimensions of quality in healthcare that Arah et al. defined as the core quality dimensions: safety, effectiveness and patient centeredness.[2] See Figure 1.


**Figure 1** Quality of Care Dimensions

## Patient safety

The assessment of patients' safety is aimed at identifying the relationships between clinical decisions and adverse outcomes, and can only be made on the basis of scientific knowledge integrated with clinical expertise. Insight into the incidence and impact of potentially unsafe situations is the first step towards improving patient safety and quality of care. There is a paucity of data on patient safety in primary care, especially in physical therapy. The safety risks that patients face in that context are different from those in hospital care due to the specific characteristics of the primary care setting.[3] Despite the low risk compared to hospital care, primary care can cause serious - though avoidable - harm to patients.[4] The large volume of contacts and procedures in primary healthcare indicate that it is important to devote attention to patient safety. For instance, one of the characteristics of primary healthcare is multi-disciplinary collaboration (*e.g.* between general practitioners and physical therapists). That kind of collaboration involves extended communication and the transfer of

information, which can be affected by various factors and give rise to unintended results. Diagnostic and intervention errors, as well as errors in communication, pose the highest risk of harming patients in primary medical care.[5]

In an earlier study of patient safety in primary care, a mix of methods was found to be necessary to identify safety incidents.[3] The retrospective patient record review is a frequently used method to retrieve the rates of safety incidents.[6] However, the validity of this method depends on the quality of record keeping.[7] Incident reporting by healthcare professionals is another method that is commonly used for the detection of latent errors, though its validity is highly dependent on the awareness and willingness of professionals to report and analyze substandard care.[8]

To increase patient safety, it is important to know which causal factors underlie unintended events. Data about the nature and causes of unintended events may help stakeholders realize how human behavior and organizational and contextual factors together cause unintended events to occur. In addition, insight into the causes of unintended events may help prioritize prevention strategies and research efforts to improve healthcare. In order to identify the causes underlying the reported unintended events, a root cause analysis tool can be used. PRISMA, an acronym for Prevention and Recovery Information System for Monitoring and Analysis, is a tool to analyze the causes of unintended events.[9,10] Unintended events are analyzed by means of causal factor trees. At the top of the causal tree, a short description of the event is placed as the starting point for the analysis. Below the top event, all direct causes are listed. By continuing to ask "why" for each event or action, all relevant causes are revealed.

The literature on patient safety in physical therapy care is scarce. Malpractice reports from different physical therapy settings in the United States have shown that the incidence of errors is low, and that such errors are mainly associated with treatment-related issues.[11] Although the majority of primary care physical therapy treatments involve little risk for the patient, most physical therapists are familiar with examples of (near) incidents. Examples include harm due to improper treatment, or due to interventions that are not risky in nature, but that can be harmful for patients with comorbidities that are contraindications to those interventions (such as spinal manipulations in patients with osteoporosis). Hazardous situations can also arise due to incorrect diagnoses, especially when red flags (signs or symptoms indicating the presence of a serious pathology) are not noticed or ignored in patients with musculoskeletal conditions. Several incidents are mentioned in the literature with respect to spinal manipulation, especially in children[12-14], but also in adults.[15] Insight into the causes of these incidents, the consequences and the severity of the harm might help to develop focused interventions to improve patient safety and the quality of physical therapy care. More insight is needed into the number and causes of incidents in the Netherlands to develop improvement strategies.

# Effectiveness

## Evidence-based practice

Evidence-based practice (EBP) is the integration of clinical expertise, patient values, and the best research evidence into the decision-making process for patient care.[16] The full integration of these three components into clinical decisions enhances the opportunity for optimal clinical outcomes and quality of life. Healthcare decisions based on sound evidence are crucial for ensuring high-quality patient care, optimal health outcomes and the quality and safety of healthcare systems. Clinical reasoning is built through the ability to think, reason and apply scientifically plausible principles in practice. A prerequisite to evidence-based practice is that healthcare professionals possess strong clinical reasoning skills. Clinical reasoning is described as the process that precedes clinical decision making. It refers to the cognitive processes associated with the clinician's examination and handling of the patient.[17] Moreover, clinical reasoning involves the individual's knowledge, self-awareness and ability to reflect on his/her approach to the thought process.

Because EBP should be patient-centered, a healthcare provider's task is to interpret the best current evidence in relation to the individual's preferences, environment, and values regarding health and well-being.[1] Clinical decision making is strongly influenced by the individual characteristics of the healthcare provider, his/her knowledge and skills, experiences, and patient perceptions.[18] Healthcare providers must rely on their clinical reasoning skills in order to practice evidence-based medicine as well as to ma ke decisions when little or no external evidence is available. Ultimately, the goal of EBP is to provide optimal clinical service to the patient on an individual basis.

## Guidelines

In order to decrease variability in clinical practice, to guide work according to the best available evidence, and to help the practitioners in their efforts to legitimize their profession in the eyes of external stakeholders, professional organizations began implementing guideline development programs since the late 1980s.[19,20] Guideline development was aimed at offering healthcare in which the healthcare provider uses the best evidence available and works in consultation with the patient to decide on care options that best suit the patient, and that reduce unwanted variation in care. This should ultimately lead to "evidence-based clinical practice." An important trigger was the increasing strain on government budgets and a pressing need for cost-effective, high quality healthcare. Clinical practice guidelines are "statements that include recommendations intended to optimize patient care that are informed by a

1 ' Patient can be used interchangeably for 'person' or 'client'

systematic review of evidence and an assessment of the benefits and harms of alternative care options."[21] They are considered an important instrument for bridging the gap between scientific evidence and actual clinical practice, and are intended to improve and support the patient-centered management and safety of the care process and health outcomes.[20,22]

The publication of guidelines does not automatically result in their use in clinical practice. Effective guideline implementation strategies involve a synergy of strategies at multiple levels.[23] Most interventions that have been applied to enhance guideline adherence had modest to moderate effects.[20,24,25] The limited success of implementation strategies is attributed to factors related to individual professionals, organizational issues, patients, and guideline quality.[25-31]

Healthcare providers may not follow guidelines if the evidence is weak, or if the recommendations do not consider patients' values and preferences.[32] For some diagnoses, it is clear what should be done, which limits the range of reasonable decisions. For other diagnoses and in cases of co-morbidity, professional uncertainty can arise regarding the proper course of action. A certain level of inter-professional variation in adherence to clinical practice guidelines is justified and even necessary due to case mix.[33] Healthcare professionals are responsible for adjusting their clinical decisions to each individual patient. Relevant arguments (*e.g.* co-morbidity, gender, patients' preferences and their personal circumstances), can justify (partial) deviation from the recommendations in the guidelines.

## Patient centeredness

Patient centeredness is defined as "healthcare that establishes a partnership among healthcare providers, patients, and their families (when appropriate) to ensure that decisions respect patients' wants, needs, and preferences, and that patients have the education and support they need to make decisions and participate in their own care."[34] In recent years, patient centeredness has gained importance in policy and practice.[35] Studies on processes and outcomes of patient-centered care generally show positive relationships between patient-centered care and outcomes, such as improved patient satisfaction, motivation, adherence to therapy, quality and outcomes of healthcare.[36-39] A Cochrane review by Lewis et al. concluded that evidence for the effects of patient-centered interventions on patient healthcare behavior or health status is mixed.[35] Elements of person-centered care include communication and relevant information, shared decision making, shared goal setting, and self-management support.[40] Several conceptual frameworks exist, describing various dimensions of patient centeredness. For example, Mead and Bower include five key dimensions in their model (biopsychosocial perspective; "patient-as-person"; sharing power and

responsibility; therapeutic alliance; and "doctor-as-person").[41] Ouwens et al. include eight domains of patient-centered care (access, follow-up, communication and respect, patient and family involvement, information, coordination, physical support and emotional and psychosocial support).[42] Patients' views on the dimensions of patient centeredness in physical therapy care include involvement in goal setting, treatment planning and outcome evaluation,[43] decision-making,[44,45] the competence and personality of the physical therapist , organization,[44] and communication.[44,46] In physical therapy practice, patient-centered care is limited due to obstacles to optimal participation and interaction. A review by Schoeb et al.[47] showed several barriers to patient participation. For one thing, they found that patients lack knowledge of what is expected from them.[47] Whereas some patients prefer to play an active role during goal setting and/or treatment, other patients prefer a passive role. It was also observed that physical therapists lack the communication skills needed to enhance patient participation and that they struggle with defining and applying the concepts of goal setting and shared responsibility.[48]

There is a growing consensus on the need for active patient involvement in order to ensure that healthcare is geared towards their needs and preferences.[49,50] Moreover, healthcare providers are becoming increasingly responsible to find out what patients want, and to support them in the goal-setting process. Goal setting is a complex interactional activity. Patients need to set goals with their physical therapists' guidance, and should aim at achieving activities that are important and meaningful to them.

Potentially, guidelines can support a patient-centered approach by guiding the healthcare provider in considering all relevant domains of the person's health, facilitating individualized and meaningful goal setting, recommending appropriate intervention strategies and using outcome measures that monitor change in the patient's health and functioning. Consideration of patients' preferences and values has been advocated in the development and implementation of clinical practice guidelines;[51-53] however, the inclusion of patients and their preferences in the process of developing guidelines is limited.[52,54] There is a potential tension between adhering to the guideline, and to optimal patient-centered care.[55] In current healthcare, patients are more involved in decision making, as patients can have different preferences concerning interventions that might influence medical decision making. Some argue that clinical practice guidelines can facilitate decision making by providing a synthesis of the research literature.[56] Guidelines can help healthcare providers in weighing the pros and cons in decision making. Evidence-based interventions should be adapted to meet individual needs and preferences, where possible. Others warn that clinical practice guidelines standardize clinical practice and limit the patient's role in decision making.[57] Moreover, evidence-based medicine may set the optimal treatment guidelines for standard patients, but it does not provide all the

answers - as has been demonstrated by many patients, conditions and situations. A translation is needed to adapt the results of clinical research, as described in guidelines, to facilitate the integration of individual patient preferences and participation in clinical decision making.

## Quality measures

Although the definitions are well defined for measuring the quality of care, the exact criteria are still a "work in progress." Quality measures may serve different purposes: for use in clinical practice at the patient level, for quality improvement among healthcare providers, for benchmark purposes to judge quality and cost performance, and to conduct comparative effectiveness studies.[58] The dimensions, as described by IOM, represent differences in quality of care perceptions among different stakeholders.[59] Professionals focus primarily on safety and effective care, whereas patients are concerned mainly with effectiveness and patient centeredness. Health care insurers and policy makers are interested in cost-effective - and timely care - at the population level. Measures of care are commonly divided into three quality aspects: structure of healthcare (organizational aspects such as equipment); process of healthcare (actual care given, such as prescribing, interaction between professionals and patients) and outcomes (the consequences of the interaction between the patient and health care provider).[60]

Guideline adherence centers mainly around process measures.[61] The quality of the care delivered can also be measured by outcome measures, which in turn, enables better assessment of patient-centered healthcare. Outcome measurements are assessments that determine changes in patients' health status, functioning, or participation in daily activities over time. The use of routine outcome measurements has several benefits. The healthcare provider receives feedback over the change in health status. That feedback can be used for communication and goal setting with the patient, and for comparison of outcomes with other health professionals.[62] Routine outcome measurement is effective in establishing diagnoses (faster and completer screening), and in monitoring treatment and communication (as it increases the frequency and effectiveness of communication).[63] Despite these benefits, outcome measurement is not routine practice. As explained in the literature, the costs of outcome measurement, misperceptions about its lack of clinical relevance, behavioral changes, and insufficient knowledge regarding the most appropriate outcome measures for different patients may be some of the obstacles preventing routine use of outcome measures in clinical practice.[64]

A patient-reported outcome measurement (PROM) of a health condition is one that comes directly from individual patients themselves. PROMs have the potential to

enhance patient-centered care and offer a promising tool to monitor and improve the quality of care.[65] PROMs data provide an added value to other measurements performed by the therapist, such as joint mobility and muscle strength, as PROMs include information regarding aspects of health problems and recovery that are important to patients. PROMs can be generic or disease-specific, and have different constructs and purposes, e.g. for measuring pain, or assessing limitations in participation or quality of life.[66] PROM measurements can be used to map complaints and limitations in the functioning of patients. That, in turn, facilitates the clinical reasoning process, goal setting and shared decision-making. The active involvement of patients in identifying their individual problems and goals increases their motivation, participation and satisfaction with their healthcare interventions.[67] Although PROMs are recommended in many different physical therapy guidelines, little is known about their actual use and the extent to which physical therapy management goals correspond with the PROMs selected.

## Quality improvement in Dutch physical therapy

In physical therapy too, clinical guidelines are considered important instruments for improving the quality of care. Guideline development has shown an impressive increase over the last decades. The principles of evidence-based practice form a core focus in the program for continuous quality improvement established by the Royal Dutch Society for Physical Therapy (KNGF). Introduced in the 1990s, this program aims at creating a continuous cycle of improvement to ensure the quality of physical therapy practice, and to promote continued professionalization.[68] The quality of care is defined in this program as care that meets a high standard, and is effective, efficient and patient-oriented.[69] The development and implementation of clinical guidelines are a key element of this program. The KNGF's strategies for implementation include mailing the guidelines to all members, publishing articles about the guidelines, and organizing symposia, conferences, and theme-based meetings in communities of practice. Despite these strategies, however, implementation remains a challenge. These methods of implementing guidelines are mainly passive interventions, and experience has shown that passive strategies are not very effective.[25,70] Adherence to the Dutch clinical guidelines in physical therapy varies considerably[71-75], and guideline-consistent behavior shows room for improvement.[76-79] The most important discrepancy between current practice and guideline recommendations in physical therapy is related to knowledge and skills, awareness of - or familiarity with - guidelines, and external factors, such as patient factors.[72,79-84]

Levels of adherence to physical therapy guidelines vary largely between physical therapists.[72,75,83,85] For example, guidelines on low back pain were followed in 53% of

all patients, but adherence levels varied from 0 to 100% among physical therapists.[72] It is generally accepted that implementation methods should be tailored to address specific barriers to guideline adherence and other features of the target group and setting.[20,86] To improve the uptake of guidelines in physical therapy, implementation strategies should focus on improving knowledge, skills, attitudes, and awareness of the need for guideline adherence.[82,83,87]

In summary, multiple elements at different levels of healthcare (e.g. organizational, policy, healthcare provider, patient) are relevant to gaining insight into the quality of care. These insights will guide the development of strategies for improving the quality of healthcare. This thesis addresses that topic by examining the following focal question: what factors can be influenced to improve the quality of healthcare in Dutch physical therapy practice?

## Outline of this thesis

This thesis aims at exploring elements of improving patient-valued quality of care, focusing on three core dimensions of quality in healthcare: safety, effectiveness and patient centeredness. The central purpose of this thesis is to identify elements that may improve the quality of physical therapy.

*Chapter two* describes the process of identifying patient safety incidents in primary care based on a patient record review of 20 allied healthcare practices (physical therapy, occupational therapy and Cesar/Mensendieck exercise therapy). The chapter also presents an overview of patient safety incidents, the causes and consequences, and the factors associated with incidents.

*Chapters three and four* report on an implementation strategy for the Dutch physical therapy guideline for low back pain. *Chapter three* describes a cluster randomized trial aimed at determining whether peer assessment is an effective strategy for improving knowledge and guideline-consistent clinical reasoning in the Dutch physical therapy guideline for low back pain. *Chapter four* addresses the perceived critical features of peer assessment that are intended to contribute to improved guideline adherence.

*Chapter five* examines the development and outcomes of an international position paper by the Allied Health Working Group at the Guidelines International Network (G-I-N) in providing recommendations on how patient centeredness might be promoted in guideline development and guideline implementation.

*Chapter six* addresses the use of patient-reported outcome measurements (PROMs) for goal setting and outcome measurements in physical therapy practices.

*Chapter seven* summarizes and discusses the main findings of the thesis, reviews strengths and limitations, and closes with an outline of the main practical and scientific implications.

This thesis concludes with a summary in both English and Dutch.

## References


1. IOM. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, D.C: Institute of Medicine, 2001.
2. Arah OA, Westert GP, Hurst J, Klazinga NS. A conceptual framework for the OECD Health Care Quality Indicators Project. Int J Qual Health Care 2006; 18 Suppl 1:5-13.
3. Wetzels R, Wolters R, van Weel C, Wensing M. Mix of methods is needed to identify adverse events in general practice: a prospective observational study. BMC Fam Pract 2008; 935.
4. Makeham MA, Mira M, Kidd MR. Lessons from the TAPS study - communication failures between hospitals and general practices. Aust Fam Physician 2008; 37(9): 735-6.
5. Wetzels R, Wolters R, van Weel C, Wensing M. Harm caused by adverse events in primary care: a clinical observational study. J Eval Clin Pract 2009; 15(2): 323-7.
6. Michel P, Quenon JL, de Sarasqueta AM, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. BMJ 2004; 328(7433): 199.
7. Rothschild JM, Landrigan CP, Cronin JW, Kaushal R, Lockley SW, Burdick E, Stone PH, Lilly CM, Katz JT, Czeisler CA, Bates DW. The Critical Care Safety Study: The incidence and nature of adverse events and serious medical errors in intensive care. Crit Care Med 2005; 33(8): 1694-700.
8. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med 2003; 18(1): 61-7.
9. van der Schaaf TW, Habraken MMP. PRISMA Method medical version a brief description. In, Eindhoven: Eindhoven University of Technology, Faculty of Technology Patient Safety Systems, 2005.
10. van Vuuren W, Shea CE, van der Schaaf TW. The development of an Incident Analysis Tool for the Medical Field. In, Eindhoven: Eindhoven University of Technology, 1997.
11. Sandstrom R. Malpractice by physical therapists: descriptive analysis of reports in the National Practitioner Data Bank public use data file, 1991-2004. J Allied Health 2007; 36(4): 201-8.
12. Vohra S, Johnston BC, Cramer K, Humphreys K. Adverse events associated with pediatric spinal manipulation: a systematic review. Pediatrics 2007; 119(1): e275-83.
13. Jacobi G, Riepert T, Kieslich M, Bohl J. [Fatal outcome during physiotherapy (Vojta's method) in a 3-month old infant. Case report and comments on manual therapy in children]. Klin Padiatr 2001; 213(2): 76-85.
14. Brand PL, Engelbert RH, Helders PJ, Offringa M. [Systematic review of the effects of therapy in infants with the KISS-syndrome (kinetic imbalance due to suboccipital strain)]. Ned Tijdschr Geneeskd 2005; 149(13): 703-7.
15. Carnes D, Mars TS, Mullinger B, Froud R, Underwood M. Adverse events and manual therapy: a systematic review. Man Ther 2010; 15(4): 355-63.
16. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312(7023): 71-2.
17. Higgs J, Jones MA, Loftus S, Christensen N. Clinical reasoning in health professions. Oxford: Butterworth-Heineman, 2008.
18. Holdar U, Wallin L, Heiwe S. Why do we do as we do? Factors influencing clinical reasoning and decision-making among physiotherapists in an acute setting. Physiother Res Int 2013; 18(4): 220-9.
19. Grol R, Wensing M. Improving patient care; the implementation of change in clinical practice. London: Elsevier, 2005.
20. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003; 362(9391): 1225-30.
21. IOM. Clinical practice guidelines we can trust. Washington, D.C.: Institute of Medicine, 2011.
22. Grol R, Cluzeau FA, Burgers JS. Clinical practice guidelines: towards better quality guidelines and increased international collaboration. Br J Cancer 2003; 89(Suppl 1): S4-8.
23. Lohr KN, Eleazer K, Mauskopf J. Health policy issues and applications for evidence-based medicine and clinical practice guidelines. Health Policy 1998; 46(1): 1-19.

24. Grimshaw JM, Thomas RE, MacLennan G, Fraser C et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6): iii-iv, 1-72.
25. Bero LA, Grilli R, Grimshaw JM, Harvey E et al. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998; 317(7156): 465-8.
26. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, Rubin HR. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999; 282(15): 1458-65.
27. Wensing M, Wollersheim H, Grol R. Organizational interventions to implement improvements in patient care: a structured review of reviews. Implement Sci 2006; 1:2.
28. Leentjens AF, Burgers JS. [What factors are important for the successful implementation of guidelines?]. Tijdschr Psychiatr 2008; 50(6): 329-35.
29. Baiardini I, Braido F, Bonini M, Compalati E, Canonica GW. Why do doctors and patients not follow guidelines? Curr Opin Allergy Clin Immunol 2009; 9(3): 228-33.
30. Chenot JF, Scherer M, Becker A, Donner-Banzhoff N et al. Acceptance and perceived barriers of implementing a guideline for managing low back in general practice. Implement Sci 2008; 3:7.
31. Solberg LI, Brekke ML, Fazio CJ, Fowles J et al. Lessons from experienced guideline implementers: attend to many factors and use multiple strategies. Jt Comm J Qual Improv 2000; 26(4): 171-88.
32. Burgers JS, Grol RP, Zaat JO, Spies TH et al. Characteristics of effective clinical guidelines for general practice. Br J Gen Pract 2003; 53(486): 15-9.
33. van der Weijden T, Legare F, Boivin A, Burgers JS et al. How to integrate individual patient values and preferences in clinical practice guidelines? A research protocol. Implement Sci 2010; 5:10.
34. Envisioning the National Health Care Quality Report. Washington, DC: Institute of Medicine, 2001.
35. Lewin SA, Skea ZC, Entwistle V, Zwarenstein M, Dick J. Interventions for providers to promote a patient-centred approach in clinical consultations. Cochrane Database Syst Rev 2001(4): CD003267.
36. Henbest RJ, Fehrsen GS. Patient-centredness: is it applicable outside the West? Its measurement and effect on outcomes. Fam Pract 1992; 9(3): 311-7.
37. Henbest RJ, Stewart M. Patient-centredness in the consultation. 2: Does it really make a difference? Fam Pract 1990; 7(1): 28-33.
38. Kinnersley P, Stott N, Peters TJ, Harvey I. The patient-centredness of consultations and outcome in primary care. Br J Gen Pract 1999; 49(446): 711-6.
39. Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ 1995; 152(9): 1423-33.
40. Robinson JH, Callister LC, Berry JA, Dearing KA. Patient-centered care and adherence: definitions and applications to improve outcomes. J Am Acad Nurse Pract 2008; 20(12): 600-7.
41. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. Soc Sci Med 2000; 51(7): 1087-110.
42. Ouwens M, Hermens R, Hulscher M, Vonk-Okhuijsen S et al. Development of indicators for patient-centred cancer care. Support Care Cancer 2010; 18(1): 121-30.
43. Payton O, Nelson C. A preliminary study of patients' perceptions of certain aspects of their physical therapy experience. Physiothery Theory Practice 1996; 1227–38.
44. Cooper K, Smith BH, Hancock E. Patient-centredness in physiotherapy from the perspective of the chronic low back pain patient. Physiotherapy 2008; 94(3): 244–52.
45. Payton O, Nelson C, Hobbs M. Physical therapy patients' perceptions of their relationships with health care professionals. Physiothery Theory Practice 1998; 14211–21.
46. Potter M, Gordon S, Hamer P. The physiotherapy experience in private practice: the patients' perspective. Aust J Physiother 2003; 49(3): 195-202.
47. Schoeb V, Burge E. Perceptions of patients and physiotherapists on patient participation: a narrative synthesis of qualitative studies. Physiother Res Int 2012; 17(2): 80-91.
48. Levack WMM, Dean SG, McPherson KM, Siegert RJ. How clinicians talk about the application of goal planning to rehabilitation for people with brain injury-variable interpretations of value and purpose. Brain Injury 2006; 20(13-14): 1439-49.
49. Florin D, Dixon J. Public involvement in health care. BMJ 2004; 328(7432): 159-61.

50. Coulter A, Ellins J. Effectiveness of strategies for informing, educating, and involving patients. BMJ 2007; 335(7609): 24-7.
51. Krahn M, Naglie G. The next step in guideline development: incorporating patient preferences. JAMA 2008; 300(4): 436-8.
52. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. Health Res Policy Syst 2006; 422.
53. McCormack JP, Loewen P. Adding "value" to clinical practice guidelines. Can Fam Physician 2007; 53(8): 1326-7.
54. Chong CA, Chen IJ, Naglie G, Krahn MD. How well do guidelines incorporate evidence on patient preferences? J Gen Intern Med 2009; 24(8): 977-82.
55. Solomon J, Raynor DK, Knapp P, Atkin K. The compatibility of prescribing guidelines and the doctor-patient partnership: a primary care mixed-methods study. Br J Gen Pract 2012; 62(597): e275-81.
56. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999; 318(7182): 527-30.
57. Greer AL, Goodwin JS, Freeman JL, Wu ZH. Bringing the patient back in. Guidelines, practice variations, and the social context of medical practice. Int J Technol Assess Health Care 2002; 18(4): 747-61.
58. Meyer GS, Nelson EC, Pryor DB, James B et al. More quality measures versus measuring what matters: a call for balance and parsimony. BMJ Qual Saf 2012; 21(11): 964-8.
59. McGlynn EA. Six challenges in measuring the quality of health care. Health Aff (Millwood) 1997; 16(3): 7-21.
60. Campbell SM, Roland MO, Buetow SA. Defining quality of care. Soc Sci Med 2000; 51(11): 1611-25.
61. Porter ME. What is value in health care? N Engl J Med 2010; 363(26): 2477-81.
62. Higginson IJ, Carr AJ. Measuring quality of life: Using quality of life measures in the clinical setting. BMJ 2001; 322(7297): 1297-300.
63. Carlier IV, Meuldijk D, Van Vliet IM, Van Fenema E et al. Routine outcome monitoring and feedback on physical or mental health status: evidence and theory. J Eval Clin Pract 2012; 18(1): 104-10.
64. Belazi D, Goldfarb NI, He H. Measuring health-related quality of life in the clinical setting. Expert Rev Pharmacoecon Outcomes Res 2002; 2(2): 109-17.
65. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract 2006; 12(5): 559-68.
66. Haywood KL. Patient-reported outcome: Measuring what matters or just another paper exercise? Musculoskeletal Care 2006; 4(2): 63-6.
67. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, Revicki DA, Symonds T, Parada A, Alonso J. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008; 17(2): 179-93.
68. Hendriks HJM, Bekkering GE, Van Ettekhoven H, Brandsma JW et al. Development and Implementation of National Practice Guidelines: A Prospect for Continuous Quality Improvement in Physiotherapy. Physiotherapy 2000; 86(10): 535–547.
69. Van der Wees PJ, Hendriks HJM, Veldhuizen HJ. Quality Assurance in the Netherlands: from development to implementationa and evaluation. Dutch J Physiotherapy 2003; 1133-6.
70. Grimshaw JM, Shirran L, Thomas R, Mowatt G et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001; 39(8 Suppl 2): II2-45.
71. Bekkering GE, van Tulder MW, Hendriks EJ, Koopmanschap MA et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys Ther 2005; 85(6): 544-55.
72. Swinkels IC, van den Ende CH, van den Bosch W, Dekker J et al. Physiotherapy management of low back pain: does practice match the Dutch guidelines? Aust J Physiother 2005; 51(1): 35-41.
73. van der Wees PJ, Hendriks EJ, Jansen MJ, van Beers H, de Bie RA et al.. Adherence to physiotherapy clinical guideline acute ankle injury and determinants of adherence: a cohort study. BMC Musculoskelet Disord 2007; 845.
74. Nijkrake MJ, Keus SH, Ewalds H, Overeem S et al. Quality indicators for physiotherapy in Parkinson's disease. Eur J Phys Rehabil Med 2009; 45(2): 239-45.

75. Rutten GM, Degen S, Hendriks EJ, Braspenning JC et al. Adherence to Clinical Practice Guidelines for Low Back Pain in Physical Therapy: Do Patients Benefit? Physl Ther 2010; 90(8): 1111-21.
76. van der Wees PJ, Jamtvedt G, Rebbeck T, de Bie RA et al. Multifaceted strategies may increase implementation of physiotherapy clinical guidelines: a systematic review. Aust J Physiother 2008; 54(4): 233-41.
77. Bekkering GE, Engers AJ, Wensing M, Hendriks HJ et al. Development of an implementation strategy for physiotherapy guidelines on low back pain. Aust J Physiother 2003; 49(3): 208-14.
78. Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. Physl Ther 2001; 81(4): 1018-28.
79. Fleuren M, Jans M. Basisvoorwaarden voor de implementatie van de KNGF-richtlijnen. [Basic requirements for implementation and Assurance of Royal Dutch Society of Physical therapy guidelines}. Leiden: TNO Kwaliteit van Leven, 2008.
80. Cote AM, Durand MJ, Tousignant M, Poitras S. Physiotherapists and use of low back pain guidelines: a qualitative study of the barriers and facilitators. J Occup Rehabil 2009; 19(1): 94-105.
81. Stevens JG, Beurskens AJ. Implementation of measurement instruments in physical therapist practice: development of a tailored strategy. Phys Ther 2010; 90(6): 953-61.
82. Harting J, Rutten GMJ, Rutten STJ, Kremers SP. A qualitative application of the diffusion of innovations theory to examine determinants of guideline adherence among physical therapists. Phys Ther 2009; 89(3): 221-32.
83. Bekkering GE, Hendriks HJM, Van Tulder MW, Knol DL et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. Qual Saf Health Care 2005; 14(2): 107-12.
84. Rutten GMJ, Harting J, Rutten STJ, Bekkering GE, Kremers SPJ. Measuring physiotherapists' guideline adherence by means of clinical vignettes: a validation study. J Eval Clin Pract 2006; 12(5): 491-500.
85. Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? Med Care 2007; 45(10): 973-80.
86. Grol R, Wensing M. What drives change? Barriers to and incentives for achieving evidence-based practice. Med J Aust 2004; 180(6 Suppl): S57-60.
87. Rutten G, Kremers S, Rutten S, Harting J. A theory-based cross-sectional survey demonstrated the important role of awareness in guideline implementation. J Clin Epidemiol 2009; 62(2): 167-76 e1.

# 2

# Patient safety in primary allied health care

## What can we learn from incidents in a Dutch exploratory cohort study?


**Published as**

Simone A. van Dulmen, Margot A.J.B. Tacken, J. Bart Staal, Sander Gaal, Michel Wensing, and Maria W.G. Nijhuis-van der Sanden. Patient safety in primary allied health care. What can we learn from incidents in a Dutch exploratory cohort study? Medical Care 2011; 49(12): 1089–1096.

## Abstract

**Background** Research on patient safety in allied healthcare is scarce. Our aim was to document patient safety in primary allied healthcare in the Netherlands and to identify factors associated with incidents.

**Design and Subject** A retrospective study of 1000 patient records in a representative sample of 20 allied healthcare practices was combined with a prospective incident-reporting study.

**Measures** All records were reviewed by trained researchers to identify patient safety incidents. The incidents were classified and analyzed, using the Prevention and Recovery Information System for Monitoring and Analysis method. Factors associated with incidents were examined in a logistic regression analysis.

**Results** In 18 out of 1000 (1.8%; 95% confidence interval: 1.0–2.6) records an incident was detected. The main causes of incidents were related to errors in clinical decisions (89%), communication with other healthcare providers (67%), and monitoring (56%). The probability of incidents was higher if more care providers had been involved and if patient records were incomplete (37% of the records). No incidents were reported in the prospective study.

**Conclusions** The absolute number of incidents was low, which could imply a low risk of harm in Dutch primary allied healthcare. Nevertheless, incompleteness of the patient records and the fact that incidents were mainly caused through human actions suggest that a focus on clinical reasoning and record keeping is needed to further enhance patient safety. Improvements in record keeping will be necessary before accurate incident reporting will be feasible and valid.

# Introduction

Patient safety has been placed high on the societal agenda,[1] although till to now research focused predominantly on hospital care.[2–4] The Netherlands has a strong primary care setting, resulting in a large volume of contacts. However, contrary to family practice,[5–7] there is hardly any concrete information on patient safety in primary allied healthcare (allied healthcare provided by professionals working in the primary care setting).[8]

A literature search on patient safety issues in allied healthcare provided only a few studies. Malpractice reports from all kinds of physical therapy (PT) settings in the United States have shown that the incidence of medical errors is low[9] and mainly associated (77%) with treatment-related issues. A survey study conducted among occupational therapists in geriatric and rehabilitation settings in the United States identified misjudgment, lack of preparation, and lack of experience as the top 3 causes of practice errors.[10] When looking more specific to primary care setting, a recent Dutch explorative study (descriptive literature review and interviews with stakeholders) concluded, that unnecessary and redundant treatment, exercise in high-risk patients, inadequate hygiene, and joint mobilizations/manipulations could be risk factors of adverse events.[8] This was affirmed in anecdotal reports with incidents after spinal manipulation in children[10–14] and adults.[15–17] So far, no studies have been found on the prevalence and types of incidents in primary allied healthcare.

One of the characteristics of primary healthcare in the Netherlands is multidisciplinary collaboration (eg, family physician and allied health therapists), which implies a necessity for alignment and clear communication. Adequate record keeping is a prerequisite for transparent, efficient, and safe care[18] and seems to decrease the risk exposure to potential malpractice.[19]

In this study, we aimed to improve our insights of safety-related issues and focused on 3 allied healthcare disciplines: PT, occupational therapy (OT), and Cesar/Mensendieck exercise therapy (ET), which is ET focusing on body awareness. The aims of this study were (1) to document the prevalence, impact, and causes of incidents and (2) to explore risk factors related to these patient safety incidents.

# Methods

## Design and setting

This study was part of a larger observational study on patient safety in primary care in the Netherlands. Identical safety studies were conducted among Dutch family practices, out-of-hours primary care centers, dental practices, and midwifery practices.[18] As a mix of methods is needed to identify incidents,[20] we choose a retrospective

patient record review as most appropriate for estimating rates of adverse events,[21] and prospective event reporting for the detection of latent errors.[22]

The same incident definition was used in both studies: "an unintended event during the care process that resulted, could have resulted or might still result in harm to the patient."[23]

## Preliminary study

In preparation for this study, a questionnaire was developed to test the feasibility of the patient record study and to trace the possible nature of incidents. We asked allied healthcare professionals about their experiences with (potentially) unsafe situations in the past years in their practice. The presented list with potentially related items on organization of care, communication, high-risk patients, diagnosis/intervention, and outcomes was based on the literature.[8,24] A link to an electronic questionnaire was sent by email to 500 allied primary care practices: a sample out of all addresses from the 3 professional organizations in the Netherlands. Therapists were asked to assess the relevance of the items and the likelihood that information could be detected in the patient records. The response rate was 262 (52%). Therapists linked (potentially) patient unsafe situations mostly to patient-related aspects, followed by lack of safe equipment as well as an inadequate history and/or missed red or yellow flags (warning signs of possible serious pathology) in the patient history. The therapists estimated that unsafe situations could adequately be identified in patient records, except for the safety of the use of equipment and material. Finally, therapists were able to point out whether they were willing to participate in the study.

## Participants

The study included 20 primary allied healthcare practices in the Netherlands. On the basis of the total number of PT, OT, and ET practices in Dutch primary care, the proportional selection was determined at 11 PT, 6 ET, and 3 OT practices. Moreover, we stratified for the degree of urbanization (half city and half country site) and only for PT on practice size [7 large ( $>$ 5 PTs) and 4 small], because OT and ET are always small practices. A secretary without knowledge about the study content appointed professionals in the practices who were willing to participate in the preliminary study ($N = 50$) from a randomly ordered list, using the grouping criteria.

## Retrospective patient record study

We randomly selected 50 records of each of the 20 practices from the appointment lists 1 to 4 months before the selection date (total 1000 records). The selection process ensured a proportional spread across the different therapists. Depending on the total number of appointed patients per therapist in practice, each "xth" record per therapist was selected. Each patient record, including correspondence and

exercise schedules, was made anonymous by the therapist, before it was reviewed. For chronic patients (treatment longer than 1 year for a single diagnosis) records were reviewed from the selection date to 1 year prior. The degree of reporting was classified based on consensus in the project group: a record was considered "good" if the information on the diagnostic and therapeutic process was complete conform the guideline reporting; "moderate" if parts of the diagnostic or therapeutic process had not been specified in the patient record and "poor" if the therapeutic or diagnostic process in the record was missing or if information on both parts was considered to be too scarce.

The review procedure was tested in 3 pilot practices (1 for each discipline), resulting in a few small alterations to the review form. Moreover, in this phase the reviewers were extensively trained in the review procedure and differences in scores were discussed.

The record review procedure consisted of 3 phases (see Fig. 1). In the first phase, reviewers assessed the sampled records. One experienced physical therapist (S.v.D.) reviewed all records using the predefined criterion list. Each practice visit she was accompanied by a researcher (M.T.) with allied healthcare background or a medical student Noortje uphoff (N.U.), all trained in the procedure. The reviewers registered on a review form potential risk factors defined by the project group based on the preliminary study and literature on incidents[8,18,24–27] (see Table 1). Risk factors included patient characteristics [sex, age $>$ 75 y, social status, high-risk patients (history of cardiovascular, lung, or cancer disease) and communication problem], intervention characteristics (intervention completed, direct access, $>$ 9 contacts, urgency for help, $>$ 1 healthcare provider involved in practice and $>$ 1 healthcare provider involved outside the practice), and record characteristic (good/moderate/poor record keeping). All patient records were completely screened by both practice visitors with special attention to the following components: adverse outcome, fall incidents, infection, or other unexpected results. Moreover, we checked not acting according to the guidelines (if available); ignoring red flags; misdiagnosis; the intensity; and content of ET (especially in high-risk patients); irregular or no assessments (objective or subjective) and no contact with referrer or other specialists. All records with only 1 score pointing into the direction of unsafe situations were copied and proceeded to phase 2. These records were assessed for potential incidents by 2 experienced physical therapists (S.v.D. and R.N.), and if necessary, consultation took place with specialized professionals. In the third phase, the same therapists (S.v.D. and R.N.) tentatively classified the causes of the incidents, the consequences and the analysis of risk factors for the cases classified as definitive incidents. Fifty records (first 10 selected cases in first 5 practices) were independently accessed (S.v.D. and R.N.) to determine the interrater reliability in phase 1. Several arrangements were made to ensure the confidentiality of the information. Patient

information was already deleted in the selected records by the therapists, therapist names were not included in the database, and reviewers and researchers signed an agreement to guarantee the confidentially of the information

## Prospective incident-reporting study

Immediately after the patient record study, all therapists of the participating practices were invited to report all unintended events during a period of 2 successive weeks. Therapists were instructed how to report all unintended events, and were stimulated to report even when they were unsure whether an event fell within that definition. They received written information on definitions of patient safety, adverse events, and incidents in advance, and more information was also available on the website.[28] If no incident occurred during the reporting period, therapists were asked to return a form in any case, so as to complete the study.

## Data analysis

Descriptive statistics and frequency tables were used to describe the population (patients and practices) and the causes of the incidents. The returned forms from the incident-reporting study were assessed to check whether these could be marked as an incident according to the stated definition.

## Causes of incidents and consequences

We applied the Prevention and Recovery Information System for Monitoring and Analysis (PRISMA) method to analyze incidents by means of causal factor trees.[29,30] It was used as a foundational component for the conceptual framework for the World Health Organization World Alliance for Patient Safety's International Classification for Patient Safety.[31–33] At the top of the causal factor tree a short description of the event is placed, as the starting point of the analysis. Below the top event, all direct causes involved are mentioned. By continuing to ask "why" for each event or action, the majority of causes are revealed. The identified root causes are classified with the Eindhoven Classification Model of PRISMA.[29,30] The Eindhoven Classification Model taxonomy distinguishes 5 main categories of causes: technical, organizational, human, patient-related, and other, which can be subdivided into subcategories. The reviewers (S.v.D. and R.N.) received extensive training in the PRISMA method.[34]

The consequences of the incidents were classified using the "severity of outcome" dimension of the International Taxonomy of Medical Errors in Primary Care.[35]

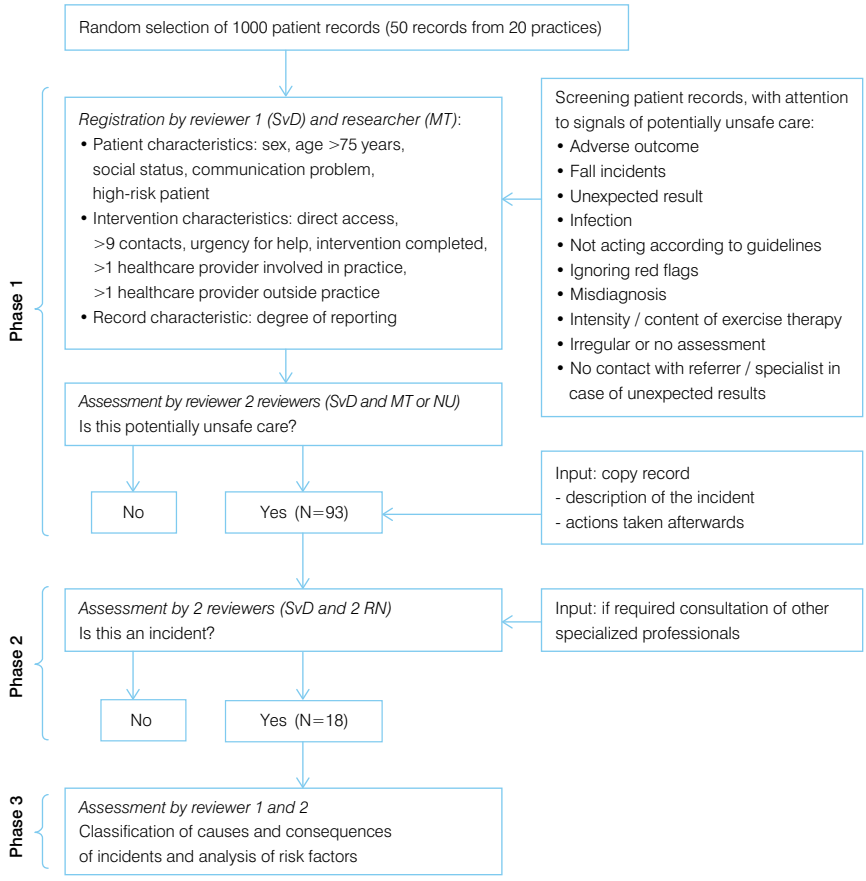


**Figure 1** Record review procedure

## Analysis of risk factors

We used logistic regression to study possible association between risk factors and the occurrence of incidents (dependent variable: "Yes an incident" vs. "No incident"). First, possible bivariate associations of all risk factors were examined with $\chi^2$ tests. Given the low number of incidents, we needed to reduce the number of variables in the model. Variables occurring in less than 20% of the incident cases were therefore excluded, assuming that the incidence would be too low to obtain reliable estimates in a logistic regression model. The remaining factors were examined in a stepwise procedure in which the risk factor with the highest *P* value was removed, whereupon the model was run again. This step was repeated until no risk factors with $P > 0.20$

had been left in the model. Results from the logistic regression model were expressed as Odds ratios (OR) including 95% confidence interval (CI). Associations were considered statistically significant if $P < 0.05$.

# Results

## Study population

We included 11 PT (550 records), 6 ET (300 records), and 3 OT practices (150 records). Nine were solo practices (45%). Most therapists (37%) were in the age category 26 to 35 years, followed by age category 46 to 55 years (29%), and 36 to 45 years (27%). On average 3.4 therapists worked per practice and 63% were women. The participating professionals in the practices and the patient records characteristics sex, age, direct access, and the number of contacts comprised a representative sample when compared with national reference data.[36,37] Only the average number of treatments per episode/year in the PT practices differed: 21 visits in this study versus 12 in the normative data due to outliers (median = 9). Table 1 presents an overview of the characteristics for the total sample and the subsample of records with incidents. Reference data of OT practices was not available.

## Prevalence of incidents in retrospective patient record study

During the first phase of the record review 93 cases with potential unsafe care were detected. It is noteworthy that only in 63% of the records the degree of reporting was "good," in 31% "moderate," and in 6% "poor." Therefore, it was sometimes difficult or even impossible to track incidents. The interrater reliability was high: (94% agreement; Cohen k: 0.83). After the assessment in phase 2, 18 cases out of 1000 (1.8%) were classified as patient safety incidents (95% CI: 1.0-2.6): 13 (2.4%) in PT records (CI: 1.1-3.6), 4 (1.3%) in ET records (CI: 0.0-2.6), and 1 (0.7%) in OT records (CI: 0.6-2.0). Of the 20 practices there were 11 practices without incidents, 4 with 1 incident [2 PT, 1 OT, and 1 ET practice(s)], 2 ET practices with 2 incidents, 2 PT practices with 3 incidents, and 1 PT practice with 4 incidents, which means that some practices were more predetermined for incidents than others.

**Table 1** Characteristics of patient records

| Risk factor | Patient records N=1000 | (% ) | Incident records N=18 | (%) |
|---|---|---|---|---|
| **Patient** | | | | |
| Sex (woman) | 622 | (62.2) | 14 | (77.8) |
| Age >75 | 99 | (9.9) | 4 | (22.2) |
| Social status (low) | 30 | (3) | 2 | (11.1) |
| Communication problem (yes) | 26 | (2.6) | 2 | (11.1) |
| High-risk patient (yes)[A] | 105 | (10.5) | 2 | (11.1) |
| **Intervention** | | | | |
| Intervention completed (no) | 540 | (54) | 13 | (72.2) |
| Direct access (yes)[B] | 194 | (19.4) | 2 | (11.1) |
| >9 contacts[C D] | 530 | (53) | 12 | (66.7) |
| High urgency for help (yes) | 40 | (4) | 1 | (5.6) |
| >1 healthcare provider involved in practice | 225 | (22.5) | 10 | (55.5) |
| >1 healthcare provider involved outside the practice | 869 | (86.9) | 18 | (100) |
| **Record** | | | | |
| Moderate/poor record keeping[E] | 366 | (36.6) | 12 | (66.7) |

[A] High-risk patient: (having had) heart problems, lung problems or cancer

[B] Direct access implies that patients came to the healthcare provider without referral

[C] In the Netherlands, Dutch National Insurance compensates 10 hours per year for occupational therapy. This means that some treatments are for 15 minutes, another for 1.5 hours. In this study occupational therapy is displayed at the number of contacts (and not at the number of hours).

[D] In the Dutch National Insurance a cut-off of more than nine treatments is used for chronic care in allied healthcare.

[E] A record was considered 'moderate' if parts of the therapeutic process or the diagnostic process had not been specified in the patient record. A record was rated as 'poor' if the therapeutic or diagnostic process in the record was missing or if information on both parts was considered to be too scarce.

**Table 2** Underlying causes of incidents according to the PRISMA-medical version (N=18)

| Category | Code | Description | Frequency | % |
|---|---|---|---|---|
| **Organizational** | | | | |
| External | O-ex | Failures beyond the control / responsibility of the organization | 1 | 5.6 |
| Knowledge transfer | OK | Failures related to communication with other care providers | 12 | 66.7 |
| Protocols | OP | Failures related to quality and availability of protocols | 0 | 0 |
| Management priorities | OM | Conflicts between production needs and safety | 0 | 0 |
| Culture | OC | Failures resulting from a collective approach and its attendant modes of behavior | 1 | 5.6 |
| **Human** | | | | |
| External | H-ex | Human failures beyond the control/ responsibility of the organization | 0 | 0 |
| Knowledge behavior | HKK | Failures in clinical decisions | 16 | 88.9 |
| Qualifications | HRQ | Incorrect fit between an individual's training and task | 0 | 0 |
| Coordination | HRC | A lack of task coordination within the organization | 4 | 22.2 |
| Verification | HRV | Failures in assessment before starting the intervention | 1 | 5.6 |
| Intervention | HRI | Failures that result from faulty task planning and execution | 8 | 44.4 |
| Monitoring | HRM | Failures in monitoring a process of patient status | 10 | 55.6 |
| Slips | HSS | Failures in performance of highly developed skills | 1 | 5.6 |
| Tripping | HST | Failures in whole body movements | 0 | 0 |
| **Patient-related** | PRF | Failures related to patient characteristics or conditions, which are beyond the control of the treatment | 1 | 5.6 |
| **Other** | X | Other | 0 | 0 |

## Prospective incident reporting study

All professionals in the 20 practices participated in the incident reporting study. In total, 7 incident report forms were returned from all 3 kinds of allied healthcare practices. However, evaluation showed no incidents according to the assigned definition. In all cases, the described events occurred in the home situation of the patient and were not directly related to the intervention and/or question for help, that is, a patient who had been involved in a car accident, or a report of a patient who fell from a wheelchair during a restaurant visit.

## Causes of incidents and consequences

Table 2 describes the classification of the causes of incidents. Almost all incidents had multiple causes. In 17 out of the 18 incidents at least 1 of the reasons was linked to human behavior. "Knowledge-based behavior" was predominantly involved (16 incidents, 89%). In all these cases there was a wrong or unclear diagnosis resulting in an incorrect intervention combined with lack of evaluation and response to unexpected results. Incidents related to "monitoring" (human rule-based monitoring) were mainly related to problems in monitoring a process or patient status (14 incidents, 56%). Incidents related to the organization of care (12 incidents, 67%) were connected to "transfer of knowledge" (organization transfer of knowledge).

Eight out of the 18 incidents did not result in actual harm. Ten events (56%) did have consequences for patients: in 2 incidents an extra intervention was needed, 2 caused emotional harm, 2 temporary physical harm, 2 patients had to be admitted to the hospital, and 2 incidents resulted in permanent harm. Table 3 describes examples of incidents with causes and consequences.

## Risk factors of incidents

Table 4 shows the relationship of the 11 prognostic variables with incidents. The variable " $>$ 1 healthcare provider involved outside the practice" was removed because it was present in nearly all cases (86.9%) and therefore not discriminative. Bivariate analyses showed that low social status, communication problems, $>$ 1 healthcare provider involved in the practice, and moderate/poor record keeping were significantly related to incidents. Owing to low frequency, 5 factors were excluded and the remaining 6 variables were included in the stepwise multivariable analyses: sex, age $>$ 75, intervention not completed, $>$ 9 contacts, $>$ 1 healthcare provider involved in practice, and moderate/poor record keeping. On the basis of the $P < 0.20$, 3 independent variables remained in the final model, of which 2 were statistically significant: treatment by more than 1 healthcare provider in the practice (OR: 3.86; CI: 1.48–10.03) and moderate/poor record keeping (OR: 3.04; CI: 1.16–8.29).

**Table 3** Examples of incidents in allied healthcare practices

| Causes of incidents* | Consequences of incidents | Examples of incidents |
|---|---|---|
| OK, HKK | Temporary harm | A patient with a hip dysplasia has to carry out an intensive exercise program. Despite increasing symptoms of decreased exercise tolerance the physical therapist did not adapt the program, resulting in a bursitis. Because of that the patient was forced to stop with the exercise program and became hindered in his functioning in daily life. |
| HKK, OK, HRI | Extra intervention necessary | A woman visited the physical therapist with a tendinitis of foot flexors with a lot of pain. Intervention included massage, rest and cooling advice. After two treatment sessions the pain suddenly increased. Patient contacted her general practitioner and further investigation showed there was a march fracture. |
| OK, HKK, HRI, HRM | Admission to the hospital | A patient has had been in a car accident, resulting in a patella luxation. After treatment in the hospital, patient got a cast for six weeks. After removing the cast the intervention by the physical therapist focused on improving mobility and stability. Patient was treated daily for two months with increased symptoms: instability, inflammation and 'giving way' of the knee. Four months later, a MRI showed a rupture of the anterior and posterior cruciate ligaments and the medial collateral ligaments. It is unclear whether the diagnosis was missed in the beginning or occurred during the exercise treatment, but it is evident that the therapist has treated extensively without progress and has not responded to a different course and 'red flags'. |
| OK, HKK, HRI, HRM, HRC | Emotional harm | Diagnosis doctor and exercise therapist: Low back pain with radiation. There is frequent use of multiple healthcare providers. Treatment exists of long-term massage therapy (48x) with no change in symptoms, no assessment and no description of communication with other healthcare providers. It is known that massage promotes a passive coping style. In addition, massage is not a defined expertise of an exercise therapist. |
| OK, HKK, HRI, HSS | Permanent harm | A man with arm complaints after a stroke lives in a nursing home and is familiar with aphasia and dementia. The physical therapist was asked to see if the arm function could be improved. Patient was treated once and the next day the arm was less functional and swollen. The therapist classified these symptoms as a normal reaction to the intervention. Two days later, by the persistence of swelling and loss of arm function, an X-ray showed an arm fracture. |

* OK=Organisation transfer of knowledge, HKK=Human knowledge based behavior, HRI=Human rule based intervention, HRM=Human rule based monitoring, HRC=Human rule based coordination, HSS=Human skill based slips

**Table 4** Bivariate and multivariable logistic regression analysis for risk factors of incidents

| Risk factor | Bivariate analysis | | Regression analysis | |
|---|---|---|---|---|
| | OR (95% CI)# | *P*-value | OR (95% CI) | *P*-value |
| **Patient** | | | | |
| Sex (woman) | 0.46 (0.15 – 1.42) | 0.169 | 0.43 (0.14-1.35) | 0.150 |
| Age >75 | 2.66 (0.86 – 8.26) | 0.077 | | |
| Social status (low) | 4.25 (0.93 – 19.41) | 0.042* | | |
| Communication problem (yes) | 4.99 (1.08 – 22.92) | 0.022* | | |
| High-risk patient (yes)[A] | 1.06 (0.24 – 4.70) | 0.932 | | |
| **Intervention** | | | | |
| Intervention completed (no) | 2.45 (0.79 – 6.34) | 0.118 | | |
| Direct access (yes)[B] | 0.51 (0.117 – 2.25) | 0.369 | | |
| >9 contacts[C D] | 1.79 (0.66 – 4.81) | 0.241 | | |
| High urgency for help (yes) | 1.42 (0.18 – 10.96) | 0.734 | | |
| >1 healthcare provider involved in practice | 4.45 (1.73 -11.43) | 0.001** | 3.86 (1.48-10.03) | 0.006** |
| **Record** | | | | |
| Moderate/poor record keeping[E] | 3.54 (1.32 – 9.53) | 0.008** | 3.04 (1.16-8.29) | 0.03* |

* Significance $P<0.05$
**Significance $P<0.01$
\# OR (95% CI): Odds Ratio 95% Confidence Interval
[A] High-risk patient: (having had) heart problems, lung problems or cancer
[B] Direct access implies that patients came to the healthcare provider without referral
[C] In the Netherlands, Dutch National Insurance compensates 10 hours per year for occupational therapy. This means that some treatments are for 15 minutes, another for 1.5 hours. In this study occupational therapy is displayed at the number of contacts (and not at the number of hours).
[D] In the Dutch National Insurance a cut-off of more than nine treatments is used for chronic care in allied healthcare.
[E] A record was considered 'moderate' if parts of the therapeutic process or the diagnostic process had not been specified in the patient record. A record was rated as 'poor' if the therapeutic or diagnostic process in the record was missing or if information on both parts was considered to be too scarce.

## Discussion

Patient safety incidents were found in 1.8% of the records in allied healthcare practices. Although the percentage suggests a low risk, we found consequences for some of the patients, which should have serious implications for the involved healthcare providers. The high number of incomplete patient records that we found may have resulted in an underestimation of incidents. This is an essential factor because we found that besides treatment by more than 1 healthcare provider in the practice, just incomplete patient records increased the risk of incidents. Almost all incidents were caused by poor knowledge, lack of monitoring and evaluation, no response to unexpected results, and inadequate communication with other healthcare providers. All these factors are also related to the availability and adequateness of information in the records.

Compared with similar studies in general practice[38] and out-of-hours primary care,[39] the frequency of incidents was low. It is difficult to compare these results with similar studies in the field of allied healthcare, as to our knowledge no similar studies have been conducted so far. At this moment, we conclude that for the individual patient this study shows that primary allied healthcare is safe. Nonetheless, more efforts should be made to avert incidents, because each incident has an impact on the individual involved and may result in high healthcare costs.

In this study, many patients were not treated according to guidelines and intervention results were not adequately monitored and evaluated. These aspects have already been signaled as risk factors for patient safety in other studies.[8,24] It is conceivable that patient care was also unsafe as a result of withholding more appropriate care and unnecessary long intervention periods.[8] Within the used definition of incidents, these cases are not detected, and therefore in future a broader evaluation could be useful. Poor record keeping seems to be a hindering factor in the transfer of care. These results correspond with findings of Mira et al,[40] who showed that the patient's perspective on adverse events is highly related to doctor rotation and good knowledge transfer. Other studies also showed the relationship between poor record keeping and adverse events[4] and an increases risk to potential malpractice.[19] To understand the implicit clinical judgment, good reporting is a key factor for the clinicians themselves and for communication between clinicians.[41] A striking finding in the preliminary study and a survey among general practitioners is that professionals themselves appoint the risk of unsafe care to patient-related characteristics,[24] rather than reflecting on their own professional behavior. In the profound analysis of causes of incidents it appeared that errors were mainly related to knowledge-based behavior, clinical reasoning, and inadequate monitoring of treatment results or reflection on what could have been expected. These findings are congruent with other studies.[10,39,42–45] The main problem was that therapists were unaware of incidents and causes, and unsafe situations were, therefore, not remarked or reported by themselves. Besides

the unawareness of riskful behavior, the absence of reported incidents in the prospective study could also be related to the short period of incident reporting. However, incident reporting is not common in allied healthcare and therapists do not focus on safety issues in daily practice, which should be improved.

## Limitations

Retrospective analyses depend on the data quality and have their limitations. First, the poor quality of the patient records hindered the detection and the assessment of incidents and the risk of unsafe situations might be larger than shown in this study. Second, in the incident analysis, we only focused on the allied healthcare patient records; patient records from other professionals were not checked. Therefore, although we acknowledged the possibility that errors were also made by other professionals in the chain of care, it was not possible to check these. Third, we only focused on incidents with possible harm for the patient and not on ineffectiveness or unnecessary care. Fourth, it was not always possible to find a causal relationship between the intervention and the (risk of) harm. Finally, positive selection bias on practice level could have lowered our estimate of the incident rate, because registration for participation was voluntary.

Despite the limitations, this study offers a first overview of the frequency, causes, and determinants of incidents in primary allied healthcare and gives a fairly good indication of the Dutch situation.

## What can be learned?

Although limited by the quality of reporting in patient records, this first exploratory study provides in-depth information about incidents in primary allied healthcare. This information is useful for the development of focused interventions to improve patient safety and quality of care.

To increase awareness of patient safety issues in allied healthcare, incident reporting should be implemented, with not only focus on riskful patient characteristics, but also on riskful professional behavior of therapists. The PRISMA analysis pointed out that clinical reasoning was crucial. Necessary scientific knowledge is formalized in guidelines, so implementation of guidelines can be an adequate tool to decrease variation in quality of care and increase safety.[46,47] However, for patient-centered interventions adequate monitoring of the treatment result and a reflexive attitude is essential. Therefore, event reporting and incident discussion as a learning tool might be an appropriate way to reconsider the knowledge and improve patient safety.48 Reflexive attitudes and clinical reasoning skills might be improved by peer review, which is based on learning in practice.[49,50] Moreover, in future, electronic patient records with integrated reminders for comparison with guidelines and evaluation will offer opportunities to monitor patient safety and quality of care.

In the next research step more attention will have to be paid to barriers and facilitators of record keeping and clinical reasoning, because these are key factors in safety. Larger practice-based observational research with mixed method strategies is necessary to get an overview of all risk factors and to find out why one practice or professional is at greater risk than another.

## References


1. Donaldson SL. An international language for patient safety. Global progress in patient safety requires classification of key concepts. Int J Qual Healthcare 2009;21:1.
2. Zegers M, de Bruijne MC, Wagner C, et al. Design of a retrospective patient record study on the occurrence of adverse events among patients in Dutch hospitals. BMC Health Serv Res 2007;7:27.
3. Hoonhout LHF, de Bruijne MC, Wagner C, et al. Direct medical costs of adverse events in Dutch hospitals. BMC Health Serv Res 2009;9:27.
4. de Bruijne MC, Zegers M, Hoonhout LHF, et al. Onbedoelde schade in Nederlandse ziekenhuizen. Dossieronderzoek van ziekenhuisopnames in 2004 [Adverse events in Dutch hospitals. Chart audit of hospital admissions in 2004]. Amsterdam/Utrecht: EMGO instituut/NIVEL; 2007.
5. Wetzels R, Wolters R, van Weel C, et al. Harm caused by adverse events in primary care: a clinical observational study. J Eval Clin Pract 2009;15:323-327.
6. Giesen P, Ferwerda R, Tijssen R, et al. Safety of telephone triage in GP cooperatives: Do triage nurses correctly estimate urgency? Qual Saf Healthcare 2007;16:181-184.
7. Sandars J, Esmail A. The frequency and nature of medical error in primary care: understanding the diversity across studies. Family Practice 2003;20:231-236.
8. Leeuw JRJ de, Veenhof C, Wagner C, et al. Patiëntveiligheid in de eerstelijnsgezondheidszorg: stand van zaken [Patient safety in primary care: State of affairs]. Utrecht: NIVEL;2008.
9. Sandstrom R. Malpractice by physical therapists: descriptive analysis of reports in the National Practitioner Data Bank public use data file, 1991-2004. J Allied Health 2007;36(4):201-208.
10. Mu K, Lohman H, Scheirton L. Occupational therapy practice errors in physical rehabilitation and geriatrics settings: A national survey study. Am J Occup Ther 2006;60(3):288-297.
11. Holla M, IJland M, van der Vliet A, et al. Overleden zuigeling na 'craniosacrale' manipulatie van hals en wervelkolom [Deceased infant after 'craniosacral' manipulation of the neck and spine]. Ned Tijdschr Geneeskd 2009;153:828-831.
12. Vohra S, Johnston B, Cramer K, et al. Adverse events associated with pediatric spinal manipulation: A systematic review. Pediatrics 2007;119:275-283.
13. Jacobi G, Riepert T, Kieslich M, et al. Über einen Todesfall während der Physiotherapie nach Vojta bei einem 3 Monate alten Säugling [Fatal outcome during physiotherapy (Vojta's method) in a 3-month old infant. Case report and comments on manual therapy in children]. Klin Padiatr 2001;213:76-85.
14. Brand PLP, Engelbert RHH, Helders PJM, et al. Systematisch literatuuronderzoek naar de effecten van behandeling bij zuigelingen met 'kopgewrichteninvloed bij storingen in de symmetrie' ('KISS-syndroom') [Systematic review of the effects of therapy in infants with the KISS-syndrome (kinetic imbalance due to suboccipital strain)]. Ned Tijdschr Geneeskd 2005; 149;703-7.
15. Carnes D, Mars TS, Mullinger B, et al. Adverse events and manual therapy: A systematic review. Man Ther 2010;355–363.
16. Kerry R, Taylor AJ, Mitchell J, et al. Cervical arterial dysfunction and manual therapy: A critical literature review to inform professional practice. Man Ther 2008;13:278-288.
17. Ernst E. Adverse effects of spinal manipulation: a systematic review. J R Soc Med 2007;100:330-338.
18. Harmsen M, Gaal S, van Dulmen S, et al. Patient safety in Dutch primary care: Study protocol. Implement Sci 2010; 5(1):50.
19. Kolber M, Lucado AM. Risk management strategies in physical therapy: documentation to avoid malpractice. Int J Healthcare Qual Assur Inc Leadersh Health Serv 2005;18(2-3):123-30.
20. Wetzels R, Wolters R, van Weel C, Wensing M: Harm caused by adverse events in primary care: a clinical observational study. J Eval Clin Pract 2009, 15:323-327.
21. Michel P, Quenon JL, de Sarasqueta AM. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. BMJ 2004;328:199.
22. Thomas EJ, Petersen LA. Measuring errors and adverse events in healthcare. J Gen Intern Med. 2003;18(1):61-7.

23. Wagner C, van der Wal G. Voor een goed begrip. Bevordering patiëntveiligheid vraagt om heldere definities [For a good understanding. Improving patient safety requires clear definitions]. Med Contact 2005;60:1888-1891.
24. Gaal S, Verstappen W, Wensing M. Patient safety in primary care: a survey of general practitioners in the Netherlands. BMC Health Serv Res 2010;10:21.
25. Peel MN. Epidemiology of falls in older age. Can J Aging 2011;30(1):7-19.
26. Hartholt KA, van Beeck EF, Polinder S, et all. Societal Consequences of Falls in the Older Population: Injuries, Healthcare Costs, and Long-Term Reduced Quality of Life. J Trauma. 2010 Nov 1. J Trauma 2011;71(3):748-53.
27. Scuffham P, Chaplin S, Legood R. Incidence and costs of unintentional falls in older people in the United Kingdom. J Epidemiol Community Health 2003;57(9):740-4.
28. Website available at: www.patientveiligheideerstelijn.nl. Accessed Oct. 10, 2010.
29. van der Schaaf TW, Habraken MMP. PRISMA method medical version. A brief description. Eindhoven: Eindhoven University of Technology, Faculty of Technology Patient Safety Systems; 2005.
30. van Vuuren W, Shea CE, van der Schaaf TW. The development of an incident analysis tool for the medical field. Eindhoven: Eindhoven University of Technology; 1997.
31. Sherman H, Castro G, Fletcher M. Towards an International Classification for Patient Safety: the conceptual framework. Int J Qual Healthcare 2009;21:2-8.
32. Runciman W, Hibbert P, Thomson R, et al. Towards an International Classification for Patient Safety: key concepts and terms. Int J Qual Healthcare 2009;21:18-26.
33. Runciman WB, Baker GR, Michel P, et al. Tracing the foundations of a conceptual framework for a patient safety ontology. Qual Saf Healthcare 2010;19:1-5.
34. VMS zorg. E-learning patient safety [in Dutch]. http://www.vmszorg.nl/4495/E-learning-patientveiligheid.html . Accessed July 16, 2009.
35. The Linnaeus-PC Collaboration: International taxonomy of medical errors in primary care - version 2. Washington DC, The Robert Graham Center; 2002.
36. Nationaal kompas Volksgezondheid [National Public Health Compass]. http://www.nationaalkompas.nl/zorg/eerstelijnszorg/paramedische-zorg. Accessed June 15, 2010.
37. Landelijke Informatievoorziening Paramedische Zorg [National Information Service for Allied Healthcare]. http://www.nivel.nl/lipz/. Accessed June 15, 2010.
38. Gaal S, Verstappen W, Wolters R, et al. Prevalence and consequences of patient safety incidents in general practice in the Netherlands: a retrospective medical record review study. Implement Sci 2011; 6:37.
39. Smits M, Huibers L, Kerssemeijer B, et al. Patient safety in out-of-hours primary care: a review of patient records. BMC Health Serv Res 2010;10:335.
40. Mira JJ, Nebo C, Lorenzo S, et al. Patient report on information given, consultation time and safety in primary care. Qual Saf Healthcare 2010;19:1-4.
41. Van Dulmen S, Schlief A, Oostendorp R, et al. Eindrapport Project Kwaliteitsindicatoren Fysiotherapie-3 (ProKwaF-3) Kwaliteit Fysiotherapeutische Verslaglegging [Project Quality indicators Physical Therapy: quality of documentation].Nijmegen: IQ healthcare; 2009.
42. Oostendorp R, Pluimers DJ, Nijhuis-van der Sanden MWG, et al. Fysiotherapeutische verslaglegging: de Achilleshiel voor Evidence-based Practice (EBP)? [Physical therapy patient documentation: the Achilles heel of evidence-based practice?] Ned Tijdschr v Fysioth 2006;116(3):56-61.
43. Bovens JAL, Pluimers DJ, Nijhuis-van der Sanden MWG, et al. Qualität der physiotherapeutischen Dokumentation in Patientenkrankengeschichte [Quality of Physical Therapy Patient Documentation]. PhysioScience 2009;5:114-123.
44. Elder NC, Dovey SM. Classification of medical errors and preventable adverse events in primary care: A synthesis of the literature. J Fam Pract 2002;51(11):927-32.
45. Wilson T, Sheikh A. Enhancing public safety in primary care. BMJ 2002;324:584-587.
46. Grol R, Wensing M, Eccles M. Improving Patient Care: The Implementation of Change in Clinical Practice. Oxford, United Kingdom: Elsevier; 2005.
47. Van der Wees PJ, Moore AP, Powers CM, Stewart A et al. Development of Clinical Guidelines in Physical Therapy: Perspective for international collaboration. Phys Ther 2011;91(10):1551-63.

48. Battles JB, Shea CE. A system of analyzing medical errors to improve GME curricula and programs. Acad Med 2001; 76 (2),125-133.
49. Hewson MG, Copeland HL, Mascha E, Arrigain S et al. Integrative medicine: implementation and evaluation of a professional development program using experiential learning and conceptual change teaching approaches.Patient Educ Couns. 2006; 62(1):5-12.
50. Gopee N. The role of peer assessment and peer review in nursing. Br J Nurs. 2001;7;10(2):115-21.

# 3

# Effectiveness of peer assessment for implementing a Dutch physical therapy low back pain guideline: cluster randomized controlled trial

**Published as**

Simone A. van Dulmen, Marjo Maas, J. Bart Staal, Geert Rutten, Henri Kiers, Maria Nijhuis-van der Sanden, Philip van der Wees. Effectiveness of peer assessment for implementing a Dutch physical therapy low back pain guideline: cluster randomized controlled trial. Physical Therapy 2014; 94(10): 1396-1409.

## Abstract

**Background** Clinical practice guidelines are considered important instruments to improve quality of care. However, success is dependent on adherence, which may be improved using peer assessment, a strategy in which professionals assess performance of their peers in a simulated setting.

**Objective** The aim of this study was to determine whether peer assessment is more effective than case-based discussions to improve knowledge and guideline-consistent clinical reasoning in the Dutch physical therapy guideline for low back pain (LBP).

**Design** A cluster randomized controlled trial was conducted.

**Setting and Participants** Ten communities of practice (CoPs) of physical therapists were cluster randomized (N=90): 6 CoPs in the peer assessment group (n=49) and 4 CoPs in the case-based discussion group (control group) (n=41).

**Intervention** Both groups participated in 4 educational sessions and used clinical patient cases. The peer assessment group reflected on performed LBP management in different roles. The control group used structured discussions.

**Measurements** Outcomes were assessed at baseline and at 6 months. The primary outcome measure was knowledge and guideline-consistent reasoning, measured with 12 performance indicators using 4 vignettes with specific guideline-related patient profiles. For each participant, the total score was calculated by adding up the percentage scores (0–100) per vignette, divided by 4. The secondary outcome measure was reflective practice, as measured by the Self-Reflection and Insight Scale (20 –100).

**Results** Vignettes were completed by 78 participants (87%). Multilevel analysis showed an increase in guideline-consistent clinical reasoning of 8.4% in the peer assessment group, whereas the control group showed a decline of 0.1% (estimated group difference=8.7%, 95% confidence interval=3.9 to 13.4). No group differences were found on self-reflection.

**Limitations** The small sample size, a short-term follow-up, and the use of vignettes as a proxy for behavior were limitations of the study.

**Conclusions** Peer assessment leads to an increase in knowledge and guideline consistent clinical reasoning.

## Introduction

Clinical practice guidelines are considered useful tools for quality improvement.[1] However, successful implementation is necessary to decrease the gap between research and current practice and to reduce costs and unwanted variability in practice.[2,3] Adherence to guideline recommendations for patients with low back pain (LBP) is associated with improved quality of care, increased activities, fewer visits, and better outcomes.[4,5] Especially for patients with LBP, new research results have accumulated over the years, requiring an update of the guideline for physical therapist management of LBP.[6] It is a challenge to implement a revised guideline when physical therapists already have a lot of experience in treating these patients. Physical therapists may have to change their behavior based on new research findings, so they need to be aware of the sometimes small but determining differences.[7]

To be successful in implementation, several barriers should be addressed, including barriers at individual, social, organizational, economic, and political levels.[8–11] Comprehensive implementation strategies are essential to increase adherence to guideline recommendations.

Research has demonstrated that guideline-consistent behavior in physical therapy shows room for improvement.[12–15] The most important discrepancy between current practice and guideline recommendations in physical therapy is related to knowledge and skills, awareness of or familiarity with guidelines, and external factors.[9,13,16,17] As regards the Dutch LBP guideline, physical therapists in the Netherlands are no exception in this respect.[18–20] A qualitative study by Harting et al.[16] identified barriers to the adoption process of guidelines, lack of practical skills, and unfavorable attitude for using guidelines. The use of measurement instruments is limited as the result of a lack of knowledge for applying, scoring, and interpreting measurement instruments.[13,21]

To improve the uptake of guidelines in physical therapy, implementation strategies should focus on improvement of knowledge, skills, attitude, and awareness of guideline adherence.[7,16,18] Small-group education and peer review are widely used methods for guideline implementation and changing professionals' performance to support critical appraisal of personal quality of care.[22–25] *Small-group education* can be defined as continuing medical education or skills training on specific subjects in a small group of peers. *Peer review* is defined as a "continuous, systematic, and critical reflection by a number of care providers, on their own and colleagues' performance, with the aim of achieving continuous improvement of the quality of care."[24(p147)] Peer review may include different methods, such as consensus development, evaluation of performance, practice visit, or peer assessment. Peer assessment is a specific form of peer review in which professionals assess (judge) the performance of their peers using relevant criteria and providing feedback.[26]

For implementing the revised LBP guideline, we hypothesized that peer assessment could be appropriate to change professional behavior. Peer assessment aims at increasing self-reflection and improving awareness of actual performance. Triggers for learning and change concern both providing and receiving feedback.[25,27–30] Several studies have been conducted to study the impact of peer assessment. Ramsey et al.[31] demonstrated that peer ratings provide a practical method to assess the performance of practicing physicians on clinical skills, humanistic qualities, and communication skills. A review by Overeem et al.[32] showed that 61% to 72% of participating physicians reported a change in their behavior using peer assessment. Similar results were found by Sargeant et al.[33] using a multisource feedback tool including peer assessment by family physicians. Case-based discussions are commonly used in postgraduate education as a strategy for implementing guidelines, stimulating reflection, and integrating scientific knowledge into clinical reasoning and decision making.[34–36] The main difference between peer assessment and case-based discussions is that peer assessment intervention focuses on assessment of performance rather than discussions.

Knowledge of and adherence to the guideline can be assessed in different ways. Self-reports are practical and inexpensive to measure clinical performance, although they may overestimate guideline adherence.[37] Using medical record review might be problematic in achieving a sufficient case mix.[38] Measurement by direct observations or using standardized patients (individuals trained to act like real patients in order to simulate a set of symptoms or problems) is expensive and timeconsuming.[32,38,39] Clinical vignettes are written patient cases that approach as much as possible the authentic context of practice. They require factual guideline knowledge as well as guideline-consistent clinical reasoning in the context of a clinical problem. Therefore, vignettes are a suitable means of assessing knowledge and evaluating guideline-consistent clinical reasoning. In assessing intentional behavior, clinical vignettes are a proxy for guideline adherence and clinical behavior.[19,38,40–44] In the present study, we compared the tailored peer assessment strategy with the case-based discussion strategy in postgraduate education. Both groups used the same clinical written patient cases. The intervention group used peer assessment in which the participants reflected on LBP management performed in 3 roles: patient, physical therapist, and assessor. Additionally, they developed and evaluated a personalized improvement plan. The control group used structured case-based group discussions with written clinical cases. The effect on knowledge and guideline-consistent reasoning was measured using clinical vignettes: descriptions of 4 patient cases with specific guideline-related patient profiles.

We hypothesized that the tailored peer assessment is more effective than case-based discussions for improving guideline knowledge, guideline-consistent clinical reasoning, and reflective practice as regular activities in postgraduate education.

The objective of our study, therefore, was to compare the peer assessment strategy with the case-based discussion strategy in postgraduate education. We used the updated Dutch LBP guideline for physical therapists[6] because of the high prevalence of this condition in clinical practice.

## Method

### Design Overview

We conducted a cluster randomized controlled trial among communities of practice (CoPs) of Dutch physical therapists from January to September 2010 to evaluate the effect of an implementation strategy on guideline knowledge and guideline-consistent clinical reasoning. Both educational programs (ie, the peer assessment and the case-based discussions) included multifaceted strategies to improve knowledge and clinical reasoning skills according to the Dutch LBP guideline for physical therapy.[6] Both educational approaches consisted of a series of four 2-hour meetings during a 6-month period. Changes in knowledge and guideline-consistent clinical reasoning were assessed with vignettes at baseline and at 6 months.

### Setting and Participants

In September 2009, all contact people of the approximately 800 existing CoPs within the professional body of physical therapists in the Netherlands (Royal Dutch Society for Physical Therapy [KNGF]) received an electronic newsletter from the secretary of the KNGF, with an invitation to choose a topic out of the approximately 30 postgraduate educational programs of the coming year. One of the programs was an educational trajectory for implementing the updated LBP guideline. Communities of practice are small groups of 5 to 15 physical therapists who share the same setting, specialization, or interests and who work together on quality improvement by choosing an educational program each year. The CoPs are broadly oriented and may include many different specializations (eg, specialization in pediatric physical therapy) and may include physical therapists working in both primary and secondary care. Communities of practice of physical therapists treating patients with LBP on a regular basis were eligible for inclusion in this educational trajectory. A meeting was organized for the interested contact people in November 2009 to provide information about the aim of the project and study procedures. After this meeting, the CoPs could decide to participate in the study.

We explored the required sample size based on an estimated important difference of at least 5% for the primary outcome measure, with an anticipated intraclass correlation coefficient (ICC) of .05 and 10% loss to follow-up. Our estimation was based on the effectiveness of audit and feedback, which generally leads to

small but potentially important improvements in professional practice, with an overall improvement of adherence to desired practice of 5%.[45] This procedure resulted in a required inclusion of 103 physical therapists in 12 clusters, which we used as the target for our study.

## Randomization and Intervention Allocation

All participants of the committed CoPs attended 1 of 2 joint meetings organized in January 2010, where the updated LBP guideline was presented and modifications in the revised guideline were explained. The participants were informed that the study would involve 2 educational strategies and that both strategies would be comparable and would require an identical time investment in 4 meetings. Randomization at the CoP level was conducted after the meetings. Communities of practice were randomized using a computerized randomization system. An independent research assistant (Angelique Schlief) who was not blinded for the allocation drew up an allocation schedule using a computerized random number generator, listed them by the number of the CoP, and subsequently assigned them to the peer assessment group or the case-based discussion group according to the allocation schedule and informed the contact people of the CoPs of the allocation by e-mail. The research assistant safeguarded the allocation codes, which were revealed only after the data analysis. The principal investigators (P.W., S.D.) did not attend the meetings with the CoPs and were blinded for the allocation of the CoPs throughout the study. After the allocation, the participating physical therapists received an electronic questionnaire to gather demographic information. Meetings were each 4 to 6 weeks, depending on the available working schedule of the physical therapists.

## Intervention: Problem-Based Peer Assessment

Peer assessment was aimed at improving guideline-consistent knowledge, clinical reasoning skills, and performance. In peer assessment, clinical performance was directly observed and evaluated by peers in a simulated setting. Participants received a peer-assessment manual in advance, which contained a description of the peer assessment procedure, a time schedule, and instructions for providing constructive feedback. Performance was assessed with a scoring sheet containing performance criteria that could be scored on a 7-point scale (1="much improvement needed," 7="no improvement needed") and some space for qualitative feedback. Performance categories addressed the diagnostic process (choice of diagnostic tests and measurement instruments, performance of clinical tests, and evaluation of outcomes) and the intervention process (choice of interventions, performance of interventions, and evaluation of outcomes).

The scoring sheet was developed and validated in another study[46] and was slightly modified to include the new guideline criteria. The peer assessment CoPs

were coached by an expert assessor (M.M.), a physical therapist with expertise in LBP and an experienced teacher. The expert assessor participated in the role of process moderator and end assessor, providing additional feedback only if necessary and when all peers had given their feedback.

During the first 2 meetings, written cases were presented, accompanied by assignments for patient role performance. Participants performed in 3 roles: physical therapist, assessor, and patient. In the physical therapist role, they were blinded for the simulation role description of the patient, so it was not known in advance what specific clinical problem was simulated. Performance in the physical therapist role included communicative skills and hands-on diagnostic and treatment skills. Choices for diagnosis and treatment were explicated by reasoning aloud. In the assessor role, participants observed the performance of their peers and provided them with oral and written feedback. In the patient role, participants simulated a clinical problem according to brief simulation guidelines. Each participant developed a personal plan for improvement, including an action plan, based on feedback and assessment of colleagues during the first 2 meetings. Each participant also completed a strengths, weaknesses, opportunities, and threats (SWOT) analysis of his or her own performance, which was evaluated and discussed with peers during the third meeting. Each physical therapist clarified his or her plan and received feedback from the other participants. In the final meeting, participants evaluated their action plan, and another session of peer assessment was scheduled. This session was identical to the first 2 meetings; however, patient cases were adapted by the assessor (M.M.) to meet the specific learning needs of the participants, such as screening of "red flags."



### Control: Case-Based Discussion

Routine case-based discussion was aimed at improving guideline-consistent knowledge and reasoning skills. Participants received a program manual that contained a structured program schedule, including a description of the case discussion procedure, a time schedule, and cases for each meeting that were given in advance. For each meeting, assignments were given to guide and evaluate the case discussion process: (1) supportive questions for unraveling the problem, (2) supportive questions for establishing a diagnosis by a physical therapist and an intervention plan, and (3) assignments to make a summary of the discussions of each meeting. After each meeting, learning results were evaluated by the group. Each participant had to explicate his or her lessons learned. During the fourth meeting, 25 written statements about the anatomical and physiological structures, etiology, diagnosis, and treatment were discussed. After this meeting, participants individually answered the statements as being true or false via an online system and received feedback on each answer from the research assistant. There was no

external coach to guide the discussion process because CoPs were familiar with this educational format.

## Outcomes and Follow-up

Outcomes were assessed at baseline and after 6 months when both groups had finalized their meetings. The primary outcome measure was knowledge of the LBP guideline and guideline-consistent reasoning, which was measured using 4 clinical vignettes developed by Rutten et al.[19] The vignettes were validated and shown to have adequate validity as a proxy measure for physical therapists' adherence to the LBP guideline. These vignettes were modified according to the updated guideline.[6] The vignettes represented 4 patient profiles: (1) a patient with acute nonspecific LBP and an expected normal recovery process; (2) a patient with subacute nonspecific LBP and an imminent delay in the recovery process (indicating that the activities and participation showed no progress during the previous 3 weeks); (3) a patient with subacute nonspecific LBP and a delayed recovery and with intervening psychosocial factors; and (4) a patient with LBP due to an underlying, serious spinal pathology (eg, inflammatory process, tumor). Profiles 1 through 3 also are presented in the LBP guideline; profile 4 is described in the guideline, but it is not labeled as a profile. Text in the vignettes was presented in separate blocks similar to the steps in the guideline. Each block was followed by questions[40] related to the assessment of patients' complaints, diagnostic activities, profile selection, the use of health outcome questionnaires, whether they would contact the referring physician, treatment objectives and strategies, expected number of treatment sessions, information and advice to be provided, planned evaluation, aftercare, and a report to the referring physician. Participants were asked to complete the questionnaires online after the joint meetings but before the start of the first group sessions and post intervention within 4 weeks after finishing the final group sessions. The score for each vignette depended on the specific guideline recommendations for specific patient profiles. The Appendix describes the changes of the revised guideline and main recommendations. Per vignette and for each step in clinical decision making, a performance indicator was used to measure guideline knowledge and guideline-consistent clinical reasoning, for a total of 12 indicators (Table 1). Performance indicators have been defined as measurable elements of practice performance that can be used to assess the quality of care.[47,48] Per vignette, for each indicator, one or more questions were formulated. Answers that matched the recommendation in the guideline were given a point, whereas answers that contravened the recommendation were given no points. For each indicator, a percentage score was calculated by dividing the actual number of correct answers by the maximum possible score and multiplying the result by 100.

**Table 1** Performance indicators to measure guideline adherence based on profiles of patients with low back pain in clinical vignettes

| | Indicator | Description |
|---|---|---|
| 1 | Red flags assessed correctly | Identification of dangerous or potentially dangerous findings in the history or examination, e.g. pain at night or unexpected body loss |
| 2 | Assessment of the patients' complaints | To assess all relevant domains in relation to a patients' health: body function, activity, participation, environmental and personal factors (according to the International Classification of Functioning, Disability and Health) |
| 3 | Correct choice of the patient profile | Patient profile determined by the course of the symptoms and factors that prevent recovery (profile 1: non-specific acute LBP and a normal recovery process; profile 2: non-specific sub-acute LBP and an imminent delay in the recovery process; profile 3: non-specific sub-acute LBP and a delayed recovery with intervening psychosocial factors; and profile 4 (not a formal profile in the guideline): LBP due to an underlying, serious spinal pathology |
| 4 | Contacting the physician in case of red flags | Contacting the physician in case of LBP due to a suspected underlying, serious spinal pathology (profile 4) |
| 5 | Choice of examination objectives related to the patient profile | Examination objectives on domains of body function, activity, participation, environmental and external factors |
| 6 | Choice of treatment objectives related to the patient profile | Treatment objectives on domains of body functions, activities, participation, environmental and personal factors |
| 7 | Choice of treatment strategies related to patient profile | Recommendations are described on treatment strategies at the start and at a later stage of the treatment |
| 8 | Number of intervention sessions | Number of sessions is limited to a maximum of three in case of acute LBP with normal course |
| 9 | Adequate information is provided | Recommendations are described on treatment strategies at the start and at a later stage of the treatment |
| 10 | Health outcome questionnaires have been applied | Measurements for diagnostic and evaluation, Numeric Rating Scale (NRS), Quebec Back Pain Disability Scale (QBPDS) or Patient-Specific Function Scale (PSFS) |
| 11 | Written report to physician | Report to the physician with information about diagnosis, intervention, number of session |
| 12 | Aftercare has been arranged | Information about what to do in case of a recurrence |

LBP = low back pain

For each vignette, the total percentage score was calculated based on the indicator scores divided by the number of indicators. In addition, a mean percentage score for overall guideline adherence was calculated by adding the 4 vignette scores and dividing the total by 4, with a score range from 0 (minimal knowledge and guideline-consistent clinical reasoning) to 100 (maximal knowledge and guideline-consistent clinical reasoning). This method is known as the *patient average method*.[49] Scores on the vignettes were calculated when at least 75% of the indicators were completed, and the overall score was calculated when at least 3 vignettes were completed (in which case, the total score was divided by 3).

The secondary outcome measure was self-reflection, measured using the Self-Reflection and Insight Scale (SRIS).[50] The SRIS is a validated instrument designed to measure the process of self-reflection and insight that is presumed as conditional to self-directed change. Reflection allows assimilation and reordering of concepts, skills, knowledge, and values into pre-existing knowledge structures and, therefore, is conditional for learning new knowledge, skills, and behavioral change.[27,51,52] The SRIS is a self-administered, 20-item closed questionnaire with a 5-point Likert scale measuring engagement and insight in reflection and need for self-reflection. The total score could range from 20 to 100, with higher scores indicating more self-reflection. The modified version of the SRIS validated by Roberts and Stark[53] was translated by 2 researchers (M.M. and P.W.), and expert validity of this version was assessed by 3 experts who judged the translation. Their comments were used to improve the Dutch version of the SRIS.

## Data Analysis

The characteristics of the participants in the 2 groups were described and tested for differences between the 2 arms using chi-square tests, unpaired *t* tests, and Mann-Whitney *U* test. Post intervention mean total scores on the 4 vignettes of each participant were included as an outcome variable in a multilevel model, and baseline scores were included as covariates. Baseline characteristics were considered confounders if they were: (1) significantly associated with the outcome variable and (2) significantly different between the groups. If both conditions were met, they were added as covariates to the multilevel model to adjust for confounding. Identical analyses were performed, with the follow-up score on the SRIS questionnaire as an outcome variable. Statistical significance was tested using 2-sided tests at a *P* value of $<.05$. To determine the associations of the score at the CoP level, we calculated the ICC[54] of the scores on the vignettes from the output of the multilevel analysis with covariance parameters included. For each indicator, we calculated mean scores at baseline and at 6 months for both groups. All statistical analyses were performed using IBM SPSS Statistics for Windows, version 20.0 (IBM Corp, Armonk, New York).

### Role of Funding Source

This study was a researcher-initiated study, primarily funded by the KNGF, with co-funding of the Radboud University Medical Center, the Scientific Institute for Quality of Healthcare, and the HAN University of Applied Sciences. The KNGF had no role in the conduct of this study, analysis or interpretation of data, or manuscript preparation.

## Results



The flow of participants is presented in Figure 1. After the invitation, 35 contact people of the CoPs attended the information meeting.

Thirteen contact people were potentially interested in participating in our study. Based on these expected CoPs, a randomization scheme was determined. Three contact people decided not to participate because of other priorities. Ten CoPs were initially cluster randomized (N=90): 6 CoPs in the peer-assessment group (n=49) and 4 CoPs in the control group (n=41).

Table 2 presents characteristics of the participants and their practices (n=78). The mean age of the participants was 42.7 years (SD=11.6), with mean of 18.7 years (SD=11.0) of practice experience, and 56% of the participants were female. The participants comprised a representative sample on age and sex compared with national reference data.[55,56] Differences in sex and number of patients with LBP treated per year were not statistically significant between the groups. The years of experience were significantly higher in the control group, but no relationship was found between the scores on the vignettes and years of experience. A significant difference was found for the proportion of manual therapists between the groups. Moreover, manual therapists had significantly higher scores on the vignettes, so this was added as covariate in the analysis.

The primary outcome measure could be analyzed for 78 (87%) of the allocated participants. After randomization, 3 participants dropped out of the study: 1 in the peer-assessment group and 2 in the case-based discussion group. Four participants in the peer-assessment group and 5 participants in the case-based discussion group had incomplete scores on the vignettes.

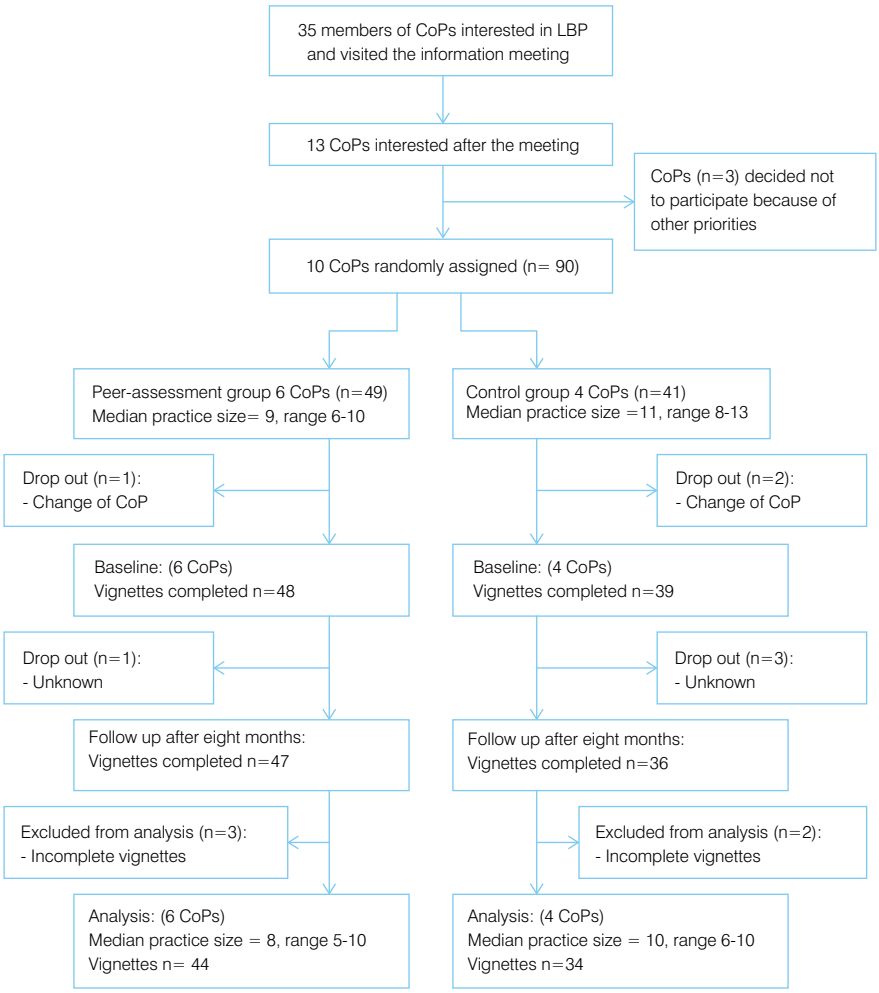


**Figure 1** Study flow diagram

Table 3 shows the mean scores for the indicators at baseline and at follow-up for the multilevel analysis. The mean increase in scores was 8.4% in the peer-assessment group, whereas scores declined 0.1% in the control group. Improvement scores on vignettes at 6 months post intervention were significantly higher in the peer-assessment group, with an estimated group difference of 8.7% ($P \leq .001$, 95% confidence interval [CI]=3.9 to 13.4). Twenty participants did not complete the SRIS, so the secondary outcome measure could be analyzed for 70 participants (78%). Mean baseline score

**Table 2** Physical therapist- and practice characteristics

| Characteristic | Peer-assessment group (n=44) | Control group (n=34) |
|---|---|---|
| Age mean (SD) | 40.4 (12.4) | 45.8 (9.9) |
| Sex (male/female) | 17/27 | 18/16 |
| Working hours per week (SD) | 32.5 (9.6) | 32.2 (10.5) |
| Treatment of patients with LBP per year | | |
| <25 | 12 | 11 |
| 25-50 | 12 | 3 |
| 50-75 | 6 | 3 |
| 76-100 | 5 | 3 |
| >100 | 10 | 13 |
| Manual therapist | 8 | 17 |
| Years of experience (SD) | 16.5 (11.9) | 21.2 (9.2) |

**Table 3** Effect of intervention on therapist knowledge and clinical reasoning and self-relfection[a]

| | Peer-assessment group | | Case-based discussion group | | Intervention effect (95% CI) | *P* |
|---|---|---|---|---|---|---|
| | Baseline | Follow up | Baseline | Follow up | | |
| *Total score on vignettes* | | | | | | .00* |
| Mean<br>SD | 63.7<br>(10.2) | 72.0<br>(11.6) | 66.8<br>(10.1) | 66.7<br>(13.1) | 8.7 (3.9 to 13.4) | |
| Range (0-100) | 44-89 | 41-98 | 47-84 | 47-87 | | |
| *SRIS score* | | | | | | .63 |
| Mean<br>SD | 74.0<br>(8.5) | 76.5<br>(9.2) | 79.9<br>(8.6) | 80.4<br>(8.6) | -0.7 (-3.5 to 2.2) | |
| Range (0-100) | 54-96 | 59-95 | 56-100 | 63-99 | | |

[a] 95% CI 95% confidence interval, SRIS=Self-Reflection and Insight Scale.

**Table 4** Mean scores and effect of intervention on therapist knowledge and clinical reasoning on the individual performance indicators

| | | Peer assessment group | | Case-based discussion group | | | |
|---|---|---|---|---|---|---|---|
| | **Indicator** | **Baseline mean (SD)** | **Follow-up mean (SD)** | **Baseline mean (SD)** | **Follow-up mean (SD)** | **Intervention effect (95% CI)** | ***P*** |
| 1 | Red flags assessed correctly | 66.3 (14.3) | 90.9 (15.3) | 70.6 (14.9) | 88.3 (15.5) | 6.7 (0.3 to 13.1) | .04* |
| 2 | Assessment of the patient's complaints | 84.8 (15.4) | 81.3 (14.7) | 86.0 (14.2) | 82.3 (15.7) | -1.5 (-8.6 to 5.7) | .69 |
| 3 | Correct choice of the patient profile | 44.3 (28.5) | 52.3 (23.0) | 51.5 (31.2) | 41.7 (23.0) | 13.1 (2.2 to 24.0) | .02* |
| 4 | Contacting the physician in case of red flags | 89.2 (12.5) | 92.6 (12.7) | 92.0 (13.4) | 89.8 (14.0) | 6.8 (1.1 to 12.4) | .02* |
| 5 | Choice of examination objectives related to the patient profile | 60.5 (10.9) | 58.9 (12.9) | 54.1 (15.3) | 52.9 (15.9) | 1.2 (-5.1 to 7.6) | .70 |
| 6 | Choice of treatment objectives related to the patient profile | 60.1 (20.4) | 67.8 (18.6) | 55.2 (19.5) | 61.2 (17.8) | 8.2 (-1.8 to 18.1) | .11 |
| 7 | Choice of treatment strategies related to patient profile | 71.6 (14.9) | 77.3 (15.4) | 67.7 (14.0) | 62.8 (19.0) | 12.4 (1.5 to 23.3) | .03* |
| 8 | Number of intervention sessions | 70.7 (24.9) | 85.8 (16.7) | 79.6 (22.2) | 85.1 (16.9) | 3.1 (-5.8 to 12.1) | .49 |
| 9 | Adequate information is provided | 43.4 (22.42 | 46.6 (21.7) | 39.5 (23.8) | 46.2 (29.3) | -0.6 (-13.6 to 12.3) | .93 |
| 10 | Health outcome questionnaires have been applied | 60.9 (23.5) | 64.1 (22.5) | 53.3 (21.0) | 56.6 (22.6) | 5.8 (-4.4 to 16.0) | .26 |
| 11 | Written report to physician | 81.6 (30.0) | 79.8 (31.8) | 86.3 (27.4) | 86.0 (30.8) | -3.9 (-18.2 to 10.3) | .58 |
| 12 | Aftercare has been arranged | 85.3 (32.0) | 94.6 (16.2) | 91.2 (26.3) | 91.7 (20.7) | 4.9 (-2.0 to 11.8) | .16 |

* $P \leq .05$

of the peer-assessment group was 74.0 points and 79.9 points of the control group. The improvement on the SRIS questionnaire was 2.5 points in the peer-assessment group and 0.5 points in the control group. The estimated difference in improvements between the 2 groups was nonsignificant (=0.69 points, *P*=.63, 95% CI=3.5 to 2.2). The ICC was <.00, indicating that the clustering effect is negligible after adjusting for covariates.

Table 4 presents the mean scores for the performance indicators at baseline and at 6 months and the results of the multilevel analysis for each indicator.



## Discussion

Our results confirmed our hypothesis that the tailored peer-assessment strategy is more effective for increasing knowledge and clinical reasoning consistent with recommendations in the LBP guideline compared with routine case-based discussions. This effect may be explained by the combination of different educational strategies: dissemination of the guideline, in-depth assessment of the guideline in a problem-solving process, assessment of performance, individualized well-timed performance feedback, and an individually tailored improvement plan. Peer assessment did not result in improved reflective practice.

The strength of the peer-assessment strategy is that participants performed different roles, which leads to a reflection on the guideline from various perspectives. In the assessor role, they had to reflect on professional qualities of colleagues using guideline recommendations as the gold standard. This reflection facilitates the ability to improve clinical skills while comparing the observed performance of colleagues with their own performance level and the guideline. In the physical therapist role, participants reflected on their own knowledge and performance using the feedback of their peers. In the patient role, they were able to reflect on the communication and perception of diagnosis and treatment from the patient's perspective. This triangle of feedback might increase reflection and awareness of individual shortcomings, which are considered key factors in guideline implementation and improvement of professional practice.[7,57] In addition, the feedback was used to develop a tailored and individualized improvement plan.

Finally, the peer assessment groups were coached by an expert assessor. By representing the gold standard, the expert assessor might have played an important role in stimulating and reinforcing the feedback and in avoiding long discussions without endpoint or consensus. It is not clear which aspects of the educational process were most attributable to the results of this study.

Intentional change of professional behavior and improved knowledge of guidelines do not necessarily lead to a concurrent change in patient outcomes. In various studies, better guideline adherence and professional behavior were not associated with improved patient outcomes.[12,58,59] Reviews focused on the effect of audit and feedback demonstrated that patient outcomes were less commonly measured and showed mixed results.[45,60] We found only one study in which a similar implementation strategy, with outcomes at the patient level, was used. In this study, peer assessment was used by general practitioners to improve care for patients with asthma and chronic obstructive pulmonary disease and showed no differences in provided care or in patients' health status.[61] Therefore, the results of our study must be interpreted with caution. Further evaluation of this strategy with appropriate designs to measure outcomes at the patient level is needed.

Although peer assessment can be a process that fosters reflection on professional quality,[62] self-reflection as measured with the SRIS did not improve in our study. For the secondary outcome measure of selfreflection, no cutoff values for clinical importance were set. We hypothesized that a minimal improvement of 5% would be of clinical importance, based on studies that assessed the effectiveness of implementation[25,63,64] and on audit and feedback.[45] There are very few studies comparing performance with selfreflection as measured with the SRIS. The SRIS was used in a course aimed at improving reflective practice of social work students and showed a significant improvement of 14.6 points on the SRIS.[65] Another study that used reflection as an approach to learning showed an improvement of 1.3 points.[66] These large variations in improvement made it difficult to reflect potential clinically important differences. Both groups in our study showed fairly high baseline scores and comparable improvement scores, indicating both interventions affected conscious reflection. The process of reflection is influenced by individual aspects and practice context.[67] Further research is necessary to identify the role of reflection in this implementation strategy and to test the validity of the SRIS in postgraduate education.

In assessing differences between the peer-assessment group and the casebased discussion group in knowledge and guideline-consistent clinical reasoning per indicator, we found lower baseline scores and significant improvements in the assessment of red flags (indicator 1), choice of the patient profile (indicator 3), contacting the referring physician in case of red flags (indicator 4), and choice of treatment strategies (indicator 7). All of these indicators include recommendations that were modified in the revised guideline, which might explain the lower baseline scores for these indicators, allowing for more improvement potential. We adjusted for the proportion of manual therapists in the multilevel analysis because there was a significant difference in the proportion of manual therapists between the 2 groups and their baseline scores on the vignettes, with higher scores for manual therapists. Manual

therapists are presumed to be familiar with topical results from clinical research on LBP, which might explain the difference in baseline scores.

Our study had several limitations; therefore, the results should be interpreted with caution. First, although peer assessment did improve knowledge and clinical reasoning consistent with recommendations in the LBP guideline, we have not demonstrated that the intervention has changed the actual behavior of physical therapists in clinical practice or resulted in better patient health outcomes. Vignettes, by construct, do not capture all important elements of care that are critical to overall patient well-being.[40] Vignettes are assumed to measure attitudes and perceptions rather than actual behavior,[68] although recent studies have demonstrated the validity of vignettes as a proxy measure of clinician performance.[19,38,40–42,44] Although the validity of vignettes used in this study was deemed acceptable,[19] other measurement instruments may be desirable.[38,44] The results of our study can be used for further analysis of using vignettes to improve knowledge and guideline-consistent clinical reasoning and to assess the relationship with clinical practice and patient outcomes.

Second, our study was conducted with a small, although representative,[55,56] self-selected sample of CoPs and physical therapists, thus threatening the external validity of the study. However, this self-selection is common in postgraduate educational interventions. Therefore, the results may be generalizable to health care professionals who are motivated to improve their quality of care and adherence to clinical practice guidelines, and, if well supervised, this method can be integrated into regular teams in primary care practices and inpatient facilities such as hospitals. We anticipate developing a course, including a training manual, for expert assessors in a national implementation program. Vignettes need to be developed for each new guideline. After allocation of 90 participants, the primary outcome was measured for 78 participants (87%), and this is a high percentage of the initial number of participants. The dropouts and participants did not differ in mean age, hours worked per week, or baseline scores on the vignettes, which suggests that the small number of dropouts did not influenced the results (data not shown). However, we did not evaluate the reasons for dropping out of the study, so it is unknown if and in which way this group affected the results.

Third, the scores on the vignettes of both groups at baseline appear to be rather high compared with results from other adherence studies.[14,19,25,69,70] This finding might be explained by the interactive educational meeting both groups attended before completing the baseline vignettes. It is known that interactive workshops can change professional practice,[71,72] which may have resulted in higher baseline scores in the current study. Furthermore, this was an updated guideline, and participants may have received educational training on the LBP guideline previously. The hypothesis that participants were familiar with the former guideline is confirmed by the lower baseline scores on the indicators with items that were modified in the revised

guideline. Moreover, registration for participation was voluntary, so selection bias of participants could have influenced the scores. They probably volunteered because they were interested in LBP or may have had a more positive attitude toward clinical guidelines, and they may have been familiar with the latest evidence in this field. Despite the high baseline scores, the effect of the intervention in the peer assessment group was 8.7%. We estimated a minimal important difference in knowledge and guideline consistent reasoning of 5%, based on improvements in professional practice using audit and feedback.[45] An update of this review showed median improvements of 4.3% for dichotomous outcomes and 1.3% for continuous outcomes.[60] Systematic reviews that assessed the effectiveness of guideline implementation showed improvements in process of care ranging from 5% to 10%.[64,73,74] The results of our study fall within this range of results.

Fourth, the peer assessment was primarily focused on knowledge and clinical reasoning of individual physical therapists, and we did not specifically address organizational and contextual barriers in the peer assessment. However, the participating physical therapists developed a personal improvement plan that allowed for addressing organizational barriers. Fifth, we did not conduct an economic evaluation of the peer-assessment program, which could be included in follow-up studies.

In conclusion, our study demonstrates that peer assessment is an effective method to improve guideline knowledge and guideline consistent clinical reasoning. Our findings are a first step toward further use of peer assessment to support the implementation of clinical guidelines and to identify areas where knowledge of guidelines should be improved. More work is needed to assess consistency of results at the patient level in clinical practice, and with professionals who are not necessarily prepared to reflect critically on their own performance. Further research should address which aspects of the educational process can be attributed to the results and assess the impact on self-reflection. Peer assessment can be integrated into CoPs of other professions as well if well prepared and supervised. Large-scale implementation can be explored by teaching expert assessor skills to group leaders within CoPs and by development of vignettes for other guidelines.

## References


1 Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004;8:iii–iv, 1–72.

2 Grol R, Wensing M. Improving Patient Care: The Implementation of Change in Clinical Practice. London, United Kingdom: Elsevier; 2005.

3 Gundersen L. The effect of clinical practice guidelines on variations in care. Ann Intern Med. 2000;133:317–318.

4 Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? Med Care. 2007;45:973–980.

5 Rutten GM, Degen S, Hendriks EJ, et al. Adherence to clinical practice guidelines for low back pain in physical therapy: do patients benefit? Phys Ther 2010;90:1111–21.

6 Staal JB, Hendriks EJM, Heijmans M, et al. Richtlijn Lage-Rugpijn voor Fysiotherapie en Manuele Therapie [Guideline Low Back Pain for Physical Therapy and Manual Therapy]. Amersfoort, the Netherlands: Royal Dutch Society for Physical Therapy; 2013.

7 Rutten G, Kremers S, Rutten S, Harting J. A theory-based cross-sectional survey demonstrated the important role of awareness in guideline implementation. J Clin Epidemiol 2009;62:167–176.e1.

8 Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines: a framework for improvement. JAMA 1999;282:1458–1465.

9 Cote AM, Durand MJ, Tousignant M, Poitras S. Physiotherapists and use of low back pain guidelines: a qualitative study of the barriers and facilitators. J Occup Rehabil 2009;19:94 –105.

10 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet. 2003; 362:1225–1230.

11 Fleuren M, Wiefferink K, Paulussen T. Determinants of innovation within health care organizations: literature review and Delphi study. Int J Qual Health Care 2004;16:107–123.

12 van der Wees PJ, Jamtvedt G, Rebbeck T, et al. Multifaceted strategies may increase implementation of physiotherapy clinical guidelines: a systematic review. Aust J Physiother 2008;54:233–241.

13 Bekkering GE, Engers AJ, Wensing M, et al. Development of an implementation strategy for physiotherapy guidelines on low back pain. Aust J Physiother 2003;49:208–214.

14 Li LC, Bombardier C. Physical therapy management of low back pain: an exploratory survey of therapist approaches. Phys Ther 2001;81:1018–1028.

15 Fleuren MA, Jans MP, van Hespen AT. Basisvoorwaarden voor de implementatie van de KNGF-Richtlijnen [basis conditions for the implementation of KNGF-guidelines]. Leiden, the Netherlands: TNO Kwaliteit van Leven; 2008.

16 Harting J, Rutten GM, Rutten ST, Kremers SP. A qualitative application of the diffusion of innovations theory to examine determinants of guideline adherence among physical therapists. Phys Ther 2009;89:221–232.

17 Stevens JG, Beurskens AJ. Implementation of measurement instruments in physical therapist practice: development of a tailored strategy. Phys Ther 2010;90:953–961.

18 Bekkering GE, Hendriks HJ, van Tulder MW, et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. Qual Saf Health Care 2005;14:107–112.

19 Rutten GM, Harting J, Rutten ST, et al. Measuring physiotherapists' guideline adherence by means of clinical vignettes: a validation study. J Eval Clin Pract 2006;12:491–500.

20 Swinkels IC, van den Ende CH, van den Bosch W, et al. Physiotherapy management of low back pain: does practice match the Dutch guidelines? Aust J Physiother 2005;51:35–41.

21 Swinkels RA, van Peppen RP, Wittink H, et al. Current use and barriers and facilitators for implementation of standardized measures in physical therapy in the Netherlands. BMC Musculoskelet Disord 2011; 12:106.

22 Speyer R, Pilz W, Van Der Kruis J, Brunings JW. Reliability and validity of student peer assessment in medical education: a systematic review. Med Teach 2011;33:e572–e585.

23 Beyer M, Gerlach FM, Flies U, et al. The development of quality circles/peer review groups as a method of quality improvement in Europe. Fam Pract 2003; 20:443–451.
24 Grol R. Quality improvement by peer review in primary care: a practical guide. Qual Health Care. 1994;3:147–152.
25 Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 2001;39(8 suppl 2):II46 –II54.
26 Falchikov N. Learning Together: Peer Tutoring in Higher Education. London: United Kingdom: Routledge Falmer; 2001.
27 Branch WT Jr, Paranjape A. Feedback and reflection: teaching methods for clinical settings. Acad Med 2002;77(12 pt 1): 1185–8.
28 Orsmond P, Merry S, Reiling K. Biology students' utilization of tutors' formative feedback: a qualitative interview study. Assess Eval High Educ 2005;30:369–386.
29 Searby M, Ewers T. An evaluation of the use of peer assessment in higher education: a case study in the School of Music, Kingston University. Assess Eval High Educ 1997;22:371–383.
30 Hanrahan SJ, Isaacs G. Assessing self and peer assessment: the students' views. High Educ Res Dev 2001;21:53–70.
31 Ramsey PG, Wenrich MD, Carline JD, et al. Use of peer ratings to evaluate physician performance. JAMA 1993;269:1655–1660.
32 Overeem K, Faber MJ, Arah OA, et al. Doctor performance assessment in daily practise—does it help doctors or not: a systematic review. Med Educ 2007;41:1039–1049.
33 Sargeant JM, Mann KV, Ferrier SN, et al. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. Acad Med.2003;78(10 suppl):S42–S44.
34 Dolmans D, Schmidt H. The advantages of problem-based curricula. Postgrad Med J 1996;72:535–538.
35 Zaher E, Ratnapalan S. Practice-based small group learning programs: systematic review. Can Fam Physician 2012;58:637–642, e310–e636.
36 Smits PB, Verbeek JH, de Buisonje CD. Problem based learning in continuing medical education: a review of controlled evaluation studies. BMJ 2002;324:153–156.
37 Adams AS, Soumerai SB, Lomas J, Ross-Degnan D. Evidence of self-report bias in assessing adherence to guidelines. Int J Qual Health Care 1999;11:187–192.
38 Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. JAMA 2000;283:1715–1722.
39 Luck J, Peabody JW. Using standardized patients to measure physicians' practice: validation study using audio recordings. BMJ 2002;325:679.
40 Peabody JW, Luck J, Glassman P, et al. Measuring the quality of physician practice by using clinical vignettes: a prospective validation study. Ann Intern Med 2004;141:771–780.
41 Hughes R, Huby M. The application of vignettes in social and nursing research. J Adv Nurs 2002;37:382–386.
42 Tiemeier H, de Vries WJ, van het Loo M, et al. Guideline adherence rates and interprofessional variation in a vignette study of depression. Qual Saf Health Care 2002;11:214–218.
43 Sandvik H. Criterion validity of responses to patient vignettes: an analysis based on management of female urinary incontinence. Fam Med 1995;27:388–392.
44 Hrisos S, Eccles MP, Francis JJ, et al. Are there valid proxy measures of clinical behaviour: a systematic review. Implement Sci. 2009;4:37.
45 Jamtvedt G, Young JM, Kristoffersen DT, et al. Audit and feedback: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2006(2):CD000259.
46 Ophery M, Maas M, de Beer J. Onderzoek naar de inhoudsvaliditeit van het performance-assessment in de hoofdfase van de bacheloropleiding fysiotherapie [Study on the content validity of the performance assessment in the main phase of the bachelor physiotherapy]. Tijdschrift voor Medisch Onderwijs 2006;25:88 -95.
47 Campbell SM, Braspenning J, Hutchinson A, Marshall MN. Research methods used in developing and applying quality indicators in primary care. BMJ. 2003;326:816–819.

48 Grol R, Cluzeau FA, Burgers JS. Clinical practice guidelines: towards better quality guidelines and increased international collaboration. Br J Cancer 2003;89(suppl 1):S4 –S8.
49 Reeves D, Campbell SM, Adams J, et al. Combining multiple indicators of clinical quality: an evaluation of different analytic approaches. Med Care. 2007;45:489–496.
50 Grant AM, Franklin J, Langford P. The Self-Reflection and Insight Scale: a new measure of private self-consciousness. Soc Behav Pers 2002;30:821–835.
51 Durgahee T. Facilitating reflection: from a sage on stage to a guide on the side. Nurse Educ Today. 1998;18:158–164.
52 Austin Z, Gregory PA, Chiu S. Use of reflection-in-action and self-assessment to promote critical thinking among pharmacy students. Am J Pharm Educ 2008; 72:48.
53 Roberts C, Stark P. Readiness for selfdirected change in professional behaviours: factorial validation of the selfreflection and insight scale. Med Educ 2008;42:1054–1063.
54 Twisk J. Applied Multilevel Analysis. New York, NY: Cambridge; 2006.
55 Deuning CM. Percentage vrouwelijke fysiotherapeuten 2010 [percentage of female physical therapists in 2010]. Available at: http://www.zorgatlas.nl/zorg/eerstelijnszorg/paramedische-zorg/percentagevrouwelijke-fysiotherapeuten#breadcrumb. Accessed November 2011.
56 Hingstman L, Kenens RJ. Cijfers uit de registratie van fysiotherapeuten: peiling 1 januari 2010 [Statistics from the Registration of Physiotherapists: Survey 1 January 2010]. Utrecht, the Netherlands: Nivel; 2011.
57 Baker R, Camosso-Stefinovic J, Gillies C, et al. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2010;(3):CD005470.
58 Rebbeck T, Maher CG, Refshauge KM. Evaluating two implementation strategies for whiplash guidelines in physiotherapy: a cluster randomised trial. Aust J Physiother 2006;52:165–174.
59 Bekkering GE, van Tulder MW, Hendriks EJ, et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys Ther 2005;85:544–555.
60 Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012;(6):CD000259.
61 Smeele IJ, Grol RP, van Schayck CP, et al. Can small group education and peer review improve care for patients with asthma/chronic obstructive pulmonary disease? Qual Health Care. 1999;8:92–98.
62 Dannefer EF, Henson LC, Bierer SB, et al. Peer assessment of professional competence. Med Educ. 2005;39:713–722.
63 Grimshaw J, Eccles M, Tetroe J. Implementing clinical guidelines: current evidence and future implications. J Contin Educ Health Prof. 2004;24(suppl 1):S31–S37.
64 Grimshaw J, Eccles M, Thomas R, et al. Toward evidence-based quality improvement: evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966–1998. J Gen Intern Med. 2006;21(suppl 2):S14–S20.
65 Chow AY, Lam DO, Leung GS, et al. Promoting reflexivity among social work students: the development and evaluation of a programme. Soc Work Educ. 2011;30:141–156.
66 Carr SE, Johnson PH. Does self reflection and insight correlate with academic performance in medical students? BMC Med Educ 2013;13:113.
67 Lowe M, Rappolt S, Jaglal S, Macdonald G. The role of reflection in implementing learning from continuing education into practice. J Contin Educ Health Prof 2007;27:143–148.
68 Gould D. Using vignettes to collect data for nursing research studies: how valid are the findings? J Clin Nurs 1996;5:207–212.
69 McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348:2635–2645.
70 Hofhuis H, Plas M, van den Ende E. Eindevaluatie van het programma Implementatie Kwaliteitsbeleid Paramedische Zorg (IKPZ): beschrijving van de kwaliteitssystemen van de paramedische beroepsgroepen in 2003 [Final Evaluation of the Program Implementation Quality in Allied Medicine (IKPZ): Description of the Quality of the Allied Professions in 2003]. Utrecht, the Netherlands: Nivel; 2004.

71 Forsetlund L, Bjorndal A, Rashidian A, et al. Continuing education meetings and workshops: effects on professional practice and health care outcomes. Cochrane Database Syst Rev. 2009;(2):CD003030.
72 Farmer AP, Legare F, Turcot L, et al. Printed educational materials: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2008;(3):CD004398.
73 Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004;8:iii–iv, 1–72.
74 Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001;286:2578–2585.

# 4

# Critical features of peer assessment of clinical performance to enhance adherence to a low back guideline for physical therapists: a mixed methods design

**Published as**

Maas MJ, van Dulmen SA, Sagasser MH, Heerkens YF, van der Vleuten CP, Nijhuis-van der Sanden MW, van der Wees PJ. Critical features of peer assessment of clinical performance to enhance adherence to a low back guideline for physical therapists: a mixed methods design. BMC Medical Education 2015; 15(1): 203.

## Abstract

**Background** Clinical practice guidelines are intended to improve the process and outcomes of patient care. However, their implementation remains a challenge. We designed an implementation strategy, based on peer assessment (PA) focusing on barriers to change in physical therapy care. A previously published randomized controlled trial showed that PA was more effective than the usual strategy "case discussion" in improving adherence to a low back pain guideline. Peer assessment aims to enhance knowledge, communication, and hands-on clinical skills consistent with guideline recommendations. Participants observed and evaluated clinical performance on the spot in a role-play simulating clinical practice. Participants performed three roles: physical therapist, assessor, and patient. This study explored the critical features of the PA program that contributed to improved guideline adherence in the perception of participants.

**Method** Dutch physical therapists working in primary care (n=49) organized in communities of practice (n=6) participated in the PA program. By unpacking the program we identified three main tasks and eleven subtasks. After the program was finished, a questionnaire was administered in which participants were asked to rank the program tasks from high to low learning value and to describe their impact on performance improvement. Overall ranking results were calculated. Additional semi-structured interviews were conducted to elaborate on the questionnaires results and were transcribed verbatim. Questionnaires comments and interview transcripts were analyzed using template analysis.

**Results** Program tasks related to performance in the therapist role were perceived to have the highest impact on learning, although task perceptions varied from challenging to threatening. Perceptions were affected by the roleplay format and the time schedule. Learning outcomes were awareness of performance, improved attitudes towards the guideline, and increased self-efficacy beliefs in managing patients with low back pain. Learning was facilitated by psychological safety and the quality of feedback.

**Conclusion** The effectiveness of PA can be attributed to the structured and performance-based design of the program. Participants showed a strong cognitive and emotional commitment to performing the physical therapist role. That might have contributed to an increased awareness of strength and weakness in clinical performance and a motivation to change routine practice.

## Background

Clinical practice guidelines are intended to optimize patient care and improve patient outcomes.[1] Guidelines are also increasingly regarded as a part of professional quality systems and policies.[2] However, the uptake of guidelines in physical therapy (PT) practice remains a challenge, despite the variety of implementation strategies that have been developed.[3–5] Professionals are hampered by a lack of commitment to the guidelines, insufficient knowledge and skills related to the guidelines, and limited social and organizational support.[6–8] In addition, a study by Rutten et al.[9] on determinants of guideline adherence showed that physical therapists (PTs) do not hold realistic perceptions of the extent to which they adhere to guideline recommendations.

The limited ability of clinicians to accurately self-assess the quality of their professional performance is not new.[10] A compelling body of research evidence shows that the development of adequate self-perception requires both internal and external information about one's professional performance, including appropriate performance standards.[11–15] There is a need for interventions containing feedback that can help to develop realistic self-perceptions of guideline adherent behavior and enhance motivation to change routine practice.

We designed an implementation strategy based on peer assessment (PA) that targets identified barriers to change for PTs in primary care.[16] We tailored an existing PA design that was shown to be effective in undergraduate PT education[17] to the context of professional PT practice and to the purpose of guideline implementation. In a previously published randomized controlled trial (Table 1), PA was shown to be more effective than the traditional "case discussion" implementation strategy.[18] We analyzed this PA program to determine the critical features of its success.

In PA professionals evaluate or are being evaluated by observing their peers in a role-play that simulates PT practice. They provide each other with performance feedback that might evoke reflection and identify areas of clinical performance that need improvement.[19,20] Personal assumptions about one's professional competence can be compared with peer views that might compensate for poor self-assessment.[13,14] Peer assessment enhances the development of a mutually accepted quality standard of performance by introducing peers to an "assessor" or "auditor" perspective.[23,26] In this respect, PA might be a an effective tool to enhance bottom up quality improvement and accountability of health care.[21,22]

Research shows that effective peer assessment practices are context-specific and culture dependent,[23,24] and these findings also apply to effective implementation strategies.[25] Thus, to enhance the generalizability of the trial results, and to allow for adequate knowledge transfer, understanding of the causal mechanisms of PA is necessary.[25–27]

**Table 1** Overview of the methods and results of a previously published trial (Van Dulmen et al. 2014)[18]

| |
|---|
| **Design** |
| A cluster-randomized controlled trial was conducted among 10 communities of practice (CoPs) of Dutch physical therapists (n=90) to compare the effectiveness of two implementation strategies: peer assessment (PA) and case discussion (CD). Both strategies aimed to improve adherence to the clinical practice guidelines for the management of patients with low back pain. The programs consisted of four meetings over a six-month period. Outcomes were measured at baseline and at 6 months follow up. |
| **Randomization and intervention allocation** |
| CoPs showing interest in the program were invited to a plenary meeting in November 2009. They were informed that the study compared two educational strategies, and that both programs required an equal amount of time and effort. All physical therapists regularly treating patients with low back pain were eligible for inclusion. Included CoPs were randomly allocated to the PA group and the CD group resulting in six CoPs for the PA program (n=49) and four CoPs for the CD program (n=41). |
| **Interventions** |
| PA is the process whereby professionals evaluate or are being evaluated by their peers and provide each other with performance feedback. The main difference between PA and CD is that in the PA approach the tasks were structured, with a focus on performance rather than discussion, and participant roles were pre-defined. In the CD approach the tasks were less structured with ample opportunity for in-depth elaboration and discussion, and participant roles were not defined. In PA and CD, participants worked on identical cases concerning problem content, but for PA these cases were adjusted to allow for performance of participants in different roles. In PA, written cases were not known in advance but were presented by a coach on the spot, simulating daily clinical practice. For CD groups, written cases were included in the program guide to allow for proper preparation, along with instructions and written questions to guide the discussion process. |
| **Outcome measures** |
| Outcomes were assessed at baseline and at six months. Primary outcome was knowledge and guideline-consistent reasoning, measured with 12 performance indicators using four vignettes that fully covered the patient profiles described in the guidelines. Changes in reflective practice were measured with the Self-Reflection and Insight Scale (Grant et al., 2002) [49] |
| **Results** |
| Vignettes were completed by 78 participants (PA group n = 44; CD group n = 34). Multilevel analysis showed an increase in guideline-consistent clinical reasoning of 8.4% in the peer-assessment groups whereas the control groups showed a decline of 0.1% (estimated group difference 8.7%; 95% CI: 3.9 to 13.4; P<0.001). No group differences were found for self-reflection. |

The aim of this study was to explore the features of the PA program that were perceived to have a powerful impact on learning and change of routine practice.

Our research question was:
Which elements of the PA program were perceived to have a strong impact on clinical performance improvement consistent with clinical guidelines, and why?

# Methods

## Study design

We conducted a mixed-methods study using questionnaires and semi-structured interviews to explore the critical features of the PA program that contributed to improved guideline adherence.

## Setting and participants

The Royal Dutch Society for Physical Therapy offers annual professional development programs for the approximately 800 communities of practice in the Netherlands. Communities of practice are small groups of 5-15 PTs who share the same setting or the same interests. The current study focused on communities of practice (n=6; 49 participants) that participated in a randomized controlled trial (Table 1) and were allocated to the PA-condition.

## The peer assessment program

The PA program was launched in February 2010 and finished in September 2010. Its design was built on a mix of theoretical constructs related to learning and professional behavior change, which were assumed to contribute to improved clinical performance.[26] Table 2 shows the theoretical framework, the underlying constructs, and the operationalization of these constructs in the PA design.

The PA program aimed to enhance clinical performance consistent with guideline recommendations including knowledge, communication, and hands-on clinical skills. Clinical performance was directly observed and evaluated by peers in a role-play that simulated clinical practice. Participants received a PA-manual in advance, containing a description of the PA procedure, a time schedule for each meeting, and guidelines for receiving and providing constructive feedback. They received a link to the updated guideline "Low back pain for physical therapy and manual therapy" (Staal et al.[28]) published by the Royal Dutch Society for Physical Therapy. Four meetings were scheduled over a period of six months. As the PTs were novices in the PA method, and no additional training was provided, the PA process

**Table 2** Theoretical framework of the PA program design

| Theory | Underlying constructs used | Operationalization of constructs |
|---|---|---|
| Social constructivist learning theory [48] | Contextual learning, collaborative learning, active participation, and knowledge construction to enhance attention, storage, and retrieval of knowledge from memory. | Presenting a variety of clinical problems that adequately reflect authentic clinical practice, accounting for the case-specifity of clinical competence.<br>Simulating the context of daily practice in a role-play accounting for the context-specifity of clinical competence.<br>Enhancing active participation of each participant by assigning pre-defined roles, and by using a performance based format. |
| Self-regulated learning theory [50,51] | Applying metacognitive strategies to guide the professional development process. | |
| | Self-assessment | Designing an improvement plan based on peer feedback. |
| | Conscious goal setting and action planning | Discussing the improvement plan with peers. |
| Situated learning theory [40,52] | Learning in the context of daily practice to bridge the gap between learning context and application context. | Delivering the program within communities of practice that share the same setting or the same interest. |
| Social cognitive learning theory [33] | Enhancing the development of self-efficacy beliefs, by: | |
| | Performing the new behavior and experiencing the consequences of that behavior (mastery experience). | Performing the new behavior individually, by reasoning aloud and demonstrating diagnostic and treatment skills relevant to the LBP guidelines. |
| | Observing the behavior of others and the consequences of that new behavior (vicarious experience). | Observing a peer's performance and providing individualized improvement feedback. |
| Stages of change theory [53] | Aligning implementation strategies to the stages of change. | Delivering the implementation program within  of practice. Peers are involved **in** the professional development process and are capable of tailoring feedback to stages of change. |

**Table 2** Continued

| Theory | Underlying constructs used | Operationalization of constructs |
|---|---|---|
| Theory of planned behaviour[34] | Changing attitudes and subjective norm toward the new behavior.<br>Enhancing the development of self-efficacy beliefs. | Introducing peers to the assessor perspective. In appraising a peers' performance, peer assessors need to develop an understanding and a mutually accepted quality standard to deliver credible performance feedback. |

was supported by a coach (MM or HE). Coaches were experienced PTs, teachers in PT education, and trained in the PA procedure. They facilitated the process of providing and receiving feedback, and they gave additional feedback when needed.

Each participant performed three roles: PT, assessor and simulated patient. In the PT role, participants completed a written assignment that contained a clinical case and brief instructions for diagnosis or treatment. Clinical cases were developed by a team of experienced PTs and guideline experts. The cases fully covered the patient profiles of LBP described in the guidelines, including red flags. PTs analyzed the clinical cases by reasoning aloud and demonstrated (hands-on) skills relevant to the clinical problem. Afterwards, they reflected on their performance. In the assessor role, peer performance was observed and assessed with a scoring sheet containing performance criteria that could be scored on a 7-point scale (1 = much improvement needed, to 7 = no improvement needed) and space for written feedback. Performance categories addressed diagnosis, treatment, and evaluation. In the patient role, participants received the clinical case along with written simulation instructions. Simulation instructions consisted of a description of the patient's complaints, including personal factors (e.g., cognitive / emotional), and contextual factors (e.g. family, work) that might be relevant to the patient's problem. Participants were instructed to improvise patient responses and provide feedback from the patient perspective.

Prior to the third session, each participant developed a personal change plan, including an action plan, based on performance feedback and self-assessment. In the third meeting, the group reviewed change plans and provided additional peer feedback. The fourth session was identical to the first two sessions, but the design of the clinical cases was tailored to participants' specific learning needs.

### Questionnaires and interviews

Prior to data collection, we unpacked the PA program and identified three main tasks and eleven subtasks that were assumed to affect guideline adherence. Task analysis was supported by guidelines described by Janssen-Noordman et al.[29] An online questionnaire was administered after completion of the PA program in which participants were asked to rank the program tasks from high to low learning value, assigning the highest rank for the most learning value and the lowest rank for the least. Subsequently, they were asked to provide written comments on the three most instructive PA task elements (Additional file 1).

Emerging questions from the questionnaires comments served as input for conducting semi-structured interviews to obtain more understanding of how the PA program affected professional development (Additional file 2). In contrast to a reductionist approach to the data by means of task analysis and task ranking, the interviews had a more holistic approach, focusing on experiences with the PA program as an integrated system. From each peer group, one PT was selected for an interview (n=6). Purposeful selection was based on average and deviant ranking results. An interview guide was designed by MM and PW addressing the three main questions that emerged from the questionnaire data:

1) What did you expect of the PA program?
2) How did you perceive the PA program, and how did it affect your daily practice?
3) In the questionnaire, you indicated that you perceived task X, Y and Z to have the strongest learning value. Can you explain why?

Selected participants were invited by e-mail, and received information about the study's purpose, procedure, the use of the data, and the focus of the interview.

The first interview was conducted by MM and PW face-to-face. The following interviews were conducted by either MM or PW using teleconferencing technology. To enhance the credibility of the results, research assistants AS and GB joined the telephone interviews, taking notes and posing additional questions when needed. Interviews of each participant, including verbal consent, were audiotaped. Interviews lasted between 45 and 90 minutes. Recordings were transcribed verbatim. An independent check on the transcripts was conducted by AS and GB.

## Data analysis

### Quantitative analysis

Ranking results were described by calculating mean, median, and sum scores for each learning task using IBM SPSS statistics 20.

### Qualitative analyses

A sample of texts from questionnaires and interview transcripts was studied and coded by MM and PW independently. The analytic process was guided by template analysis that combines a-priori codes with emerging codes.[30] The PA program as a whole and its learning tasks and subtasks served as a-priori codes. Additional codes were defined during the analytic process when these seemed relevant regarding the research question. Codes were compared, and some codes were merged into higher-order codes. PW and MM discussed a codebook codes (tasks and subtasks) and emerging themes from the data.[31] Two independent researchers SD (health scientist and PT) and MS (educational scientist) evaluated the analysis process and outcomes. They were not involved in the design or delivery of the PA program. Disagreements were discussed until consensus was reached and we finally agreed that the matrix fully fitted the data.

### Ethical aspects

This project received approval of the medical ethical committee of Radboud University Medical Center. All participants volunteered to participate and gave their informed consent. We adhered to the RATS guidelines for qualitative research.[32]



## Results

In total, 44 PTs have finished the program. Table 3 shows an overview of the participants' characteristics. Two PTs did not fully complete the ranking procedure and were excluded from quantitative analyses (response rate = 86%). All PTs invited for additional interviews (n=6) agreed to participate.

### Results quantitative analysis

Ranking results showed that participants committed the most to subtasks related to task performance in the PT role. Receiving peer feedback was perceived as the most valuable element, followed by receiving external coach feedback, performing the clinical task individually, and receiving simulated patient feedback. Participants varied widely in their preferences for learning in the PT role, but agreed on the superior value of receiving peer feedback. Table 4 shows an overview of the results.

**Table 3** Peer assessment group characteristics

| Physical therapist characteristics | N = 44 |
|---|---|
| Age mean (SD) | 40.4 (12.4) |
| Sex (male/female) | 17/27 |
| Working hours per week (SD) | 32.5 (9.6) |
| Treatment of patients with LBP per year | |
| <25 | 12 |
| 25-50 | 12 |
| 50-75 | 6 |
| 76-100 | 5 |
| >100 | 10 |
| Manual therapist | 8 |
| Years of experience (SD) | 16.5 (11.9) |

**Table 4** Results quantitative analysis

| Tasks | Subtasks | Mean | Median | Range | Sum |
|---|---|---|---|---|---|
| Study manual | Study PA procedure and guidelines | 5.09 | 6.0 | 10 | 195 |
| Perform in PT role | Perform clinical task individually | 8.05 | 9.0 | 10 | 322 |
| | Receive peer feedback | 9.75 | 10.0 | 6 | 389 |
| | Receive external coach feedback | 8.48 | 9.0 | 10 | 331 |
| | Receive simulated patient feedback | 6.84 | 7.0 | 9 | 253 |
| | Receive written feedback and scores | 2.91 | 2.0 | 9 | 102 |
| Perform in assessor role | Observe peer performance | 6.46 | 6.0 | 9 | 252 |
| | Provide oral feedback | 5.75 | 5.5 | 9 | 230 |
| | Provide written feedback and scores | 2.58 | 2.0 | 4 | 44 |
| Design change plan | Design and discuss change plan | 6.38 | 6.0 | 10 | 249 |
| Perform in patient role | Simulate patient problem | 3.26 | 3.0 | 7 | 98 |

## Results qualitative analysis

Five themes emerged from the analysis of the questionnaires comments and the additional interview transcripts. These themes were related to the PA program either as a whole, or related to its specific learning tasks and subtasks: a) general perceptions of the PA program, b) determinants of PA affecting perceptions, c) facilitators for learning, d) learning activities, and e) learning outcomes.

We summarized the results by creating a matrix that crossed a-priori categories (program tasks and subtasks) with emerging themes, leaving empty fields where data were not available (Table 5). Program tasks and subtasks in the matrix follow the build-up of the PA program. In the next section, we first discuss the general perceptions of the PA program, determinants of PA affecting these perceptions, and the general outcomes. Second, we discuss the subtasks by following the matrix, including their related learning activities, outcomes, and facilitators for learning. Although we did not explicitly ask participants to comment on tasks that were perceived as less instructive, they often did so spontaneously:

> *"Receiving feedback from your colleagues provides new insights. You learn from the mistakes you make, or how you can handle them better. I assigned the lowest ranks to 'receiving and providing scores' because I think that scores add nothing to the learning process.*
> *Moreover not all aspects of performance can be expressed in scores and scores are not objective" (Q-P8).*



We limit the discussion to comments on the most instructive subtasks. Participants' quotes are coded by information source (Questionnaire = Q; Interview Transcript = IT) and by participant number (P1–P42) (Table 5).

## The PA program as a whole

### *General perceptions*

Participants were generally satisfied with the program. They reported that the mix of written cases adequately reflected the problems encountered in daily practice, however, the PA format was new, and was perceived with mixed feelings. Physical therapists were not used to exposing their professional performance for group review. Some participants appraised the PA program as challenging, providing an excellent opportunity to receive performance feedback; others were reluctant to expose their professional competence, triggered by feelings of performance anxiety.

Specific task features (time schedule and role-play format) affected perceived learning opportunities and threats. Participants, who appreciated the task structure, reported that PA allowed them to solve a considerable number of clinical cases in a relatively short time and trained them to be concrete and concise in reasoning aloud in the PT role as well as in the assessor role.

> *"The strongest feature of PA was the structure of the meetings. The system of PA was interesting… for example, I appreciated that repeating feedback that was provided by someone else, was not allowed. It's useless to repeat advice."(IT-P41)*

**Table 5** Summary of results qualitative analysis

| | PA Program tasks and subtasks | Perceptions of the PA program | Determinants of PA affecting perceptions | Facilitators for learning and change | Learning Processes | Learning Outcomes |
|---|---|---|---|---|---|---|
| | PA Program | | | | | Change in attitudes toward guidelines.<br>Awareness of professional limitations. |
| 1 | Study manual | | | | | Update of knowledge. |
| 2 | Perform task in PT role | Fear to expose professional competence.<br>Challenge of obtaining performance feedback. | Tight time schedule.<br>Role play format. | Training in the PT role.<br>Group safety | Uncovers weakness.<br>Reinforces strength.<br>Stimulates reasoning aloud, self-assessment and critical reflection. | Awareness of gaps in professional performance.<br><br>Improved self-confidence in arguing for choices.<br><br>Improved self-efficacy beliefs in managing LBP* patients. |
| 3 | Receive peer feedback | | | Peer feedback is concrete, concise, critical and personal.<br>Varied group composition | Reveals strength and weakness in clinical performance.<br>Shows improvement areas. Reveals new reasoning perspectives and performance alternatives.<br>Stimulates self-assessment and critical reflection. | |
| 4 | Receive simulated patient feedback | | | | Reveals how interventions are perceived from the patient perspective. | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 5 | Receive external coach feedback | | | External coach poses challenging questions, guides the PA process, facilitates giving and receiving feedback, provides non-judgmental, concise feedback, monitors the time schedule, maintains group safety. | Reveals new reasoning perspectives and performance alternatives. Stimulates self-assessment and critical reflection. | |
| 6 | Receive written feedback and scores | | | | Stimulates self-assessment and critical reflection. | |
| 7 | Observe peer performance | | | Modeling peer performance. | Reveals new reasoning perspectives and performance alternatives. | Improved self-confidence in managing LBP patients. |
| 8 | Provide oral feedback | | | Training in the assessor role. | Triggers being concrete and concise in reasoning aloud. Elicits discussion over criteria. | Shared quality standards of performance. |
| 9 | Provide written feedback and scores | | | | | |
| 10 | Design change plan | | | | Guides improvement process. | |
| 11 | Perform task in Simulated patient role | | | | | |

*LBP: low back pain

Participants who criticized the task structure perceived the timetable as stressful, and as a barrier to in-depth case discussion.

> *"Yes, time pressure was a weakness of PA…. sometimes the performance evaluation raised questions which could not be addressed in-depth, because you had to skip to a new problem. I would prefer to perhaps discuss fewer cases more extensively".(IT-P18)*

From the perspective of the assessor, the role-play was appreciated because it allowed implicit behaviors to become explicit. From the perspective of the assessed, the role-play was critically appraised. Some participants believed that it poorly reflected their authentic professional behaviors, and that they underperformed in the PA context.

> *"It was hard to perform a clinical examination or treatment in this setting; partly, because the patient is a colleague. It is not like in your own working room. In addition, you consciously think about the decisions you make, because your steps will be evaluated." (Q-P8)*

### *General learning outcomes*

The PA program resulted in distinct levels of self-reported behavioral change. Although participants studied the updated guidelines prior to the program and were tested on their knowledge with clinical vignettes, they reported that applying knowledge in the context of PA increased their understanding of the guidelines, and facilitated their use in clinical practice.

> *"Yes, you want to work according to the guidelines. Therefore, you need to master them…I realized that I in fact did not fully understand the guidelines for low back pain. I knew vaguely what the content was, but not exactly. I think I have obtained a better understanding of the classification system of patient profiles, and therefore I apply them more frequently in my work." (IT-P18)*

Participants noticed that working with the guidelines in the context of the PA program changed their attitudes towards the guidelines. In their view, guidelines are often considered as too theoretical and of limited applicability in daily practice.

> *"I also noticed that some colleagues perceived the guidelines as less annoying or boring." (IT-P18)*

Although participants did not explicitly report changes in their management of patient problems, they did report changes in their professional identity and awareness of the limitations of their profession.

> *"What clearly emerged from the cases we discussed in the PA program was that as a PT we like to help people and it remains questionable if that is always justified? We somehow suffer from an irrepressible desire to help….we're inclined to always give care, whereas in some cases restraint would be better." (IT-P14)*

## Performing the PT role

### *Performing the clinical task individually*

Although some participants initially felt reluctant to move out of their "comfort zone", they considered exposure of their routine practice as a necessity for quality improvement. They pointed out that the four PA sessions allowed them to cope with anxiety triggers by training in the PT role.

> *"Yes, but you need to push yourself sometimes. I mean…I think it's threatening, it's not pleasant at all…… but I also know that it is important to bare your buttocks, and look where you go wrong. No pain no gain, that's a bit of the rationale."(IT-P15)*

Performance in the PT role necessitated reasoning aloud, triggered underpinning clinical decisions, and stimulated the transfer of research evidence to the context of a particular clinical problem. Participants explained that arguing aloud resulted in improved self-confidence in decision-making. They became more aware of their strengths and weaknesses, either by "reflection in action" or by ' "reflection on action". Exposing professional performance in the PT role was facilitated by perceived group safety.

> *"Your colleagues are the people who know you well and who know what your strengths and your weaknesses are. So they may well shoot at you."(IT-P18)*

### *Receiving peer feedback*

Although PTs organized in communities of practice discuss clinical cases on a regular basis, they do not have a culture of asking for and providing performance feedback.

The opportunity to receive peer feedback was therefore embraced. Participants felt strengthened in areas of clinical performance they mastered, and felt challenged to appraise areas that needed improvement.

*"Receiving peer feedback clearly revealed my strengths and weaknesses. I immediately understood what I needed to work on. And because my strengths were noticed, it was easier to face my weaknesses." (Q-P7)*

Learning from peer feedback was facilitated by its quality. Participants preferred personalized feedback, that showed involvement with their development process and their personal learning needs, but feedback should also focused.

*"I don't mind when someone criticizes me….of course I like to know if I'm doing right, but I'd rather know what I can improve, and how."(IT-P18)*

Another facilitating factor was the heterogeneity in group composition. Differences in age and specialization allowed for different approaches to health problems and different models of reasoning. Because feedback providers were encouraged to clarify improvement feedback with clear examples of desired behavior, they discovered new reasoning perspectives and performance alternatives.

*"For example, we have a specialist in haptonomy in our team, and he brings in new perspectives on health problems….I profit from his views in my daily practice. For example, I try to keep the global overview instead of focusing on a single vertebra. As a manual therapist I tend to focus on the details and lose the whole picture." (IT-P14)*

### *Receiving external coach feedback*

In contrast to peer feedback, participants attributed the value of coach feedback to its objectivity, conciseness, and perceptiveness, rather than to its involvement with individual peers.

*"Well, the coach had an objective approach. The feedback was very practical and well summarized. Nothing more, nothing less and because the coach was new, feedback was perceived to be more objective. I also noticed that the coach was able to discover strengths in all participants." (IT-P2)*

However, from the PT-role perspective, the presence of the coach raised performance stress in some cases.

*"We also needed to get used to her coach. At least, that applied to me. You need to feel a kind of safety with each other to show openly what you think and what you do. We share this safety in our group, and that allows us not to mince words. But with a strange person here, the threshold is higher, at least in my opinion".(IT-P1)*

Facilitating behaviors from the coach included posing critical questions rather than giving straightforward answers, fostering a safe learning environment, monitoring the structure and the time-schedule of the PA process, facilitating peer feedback delivery, and strengthening group learning. Participants rejected too much interference of the coach and judgmental coach feedback.

Receiving simulated patient feedback Participants varied in their appreciation of simulated patient feedback, referring to the limitations of role-play. Despite its limitations, participants valued the different perspective of patient feedback.

> *"While performing the assignment, I noticed that I was not always providing clear information…I previously never thought about that … I have learned now that I need to communicate more carefully, for instance when giving bad news." (Q-P12).*



### *Performing the assessor role*

Observing a peer's performance participants reported that the role of assessor allowed them to mirror and model the observed performance to their own intended performance.

> *"I found observing a peer's performance very instructive because you often imagine how you would handle the situation. When you see how your colleague deals with a problem, you critically reflect on your own choices." (Q-P19)*

Appraising the performance of a peer was not a common practice. Participants would rather discuss than assess the observed behaviors. Giving instructive feedback (according to the feedback guidelines) was perceived as difficult. It required clear reasoning strategies, arguing for quality standards of performance, and the courage to be critical.

> *"Your own feedback should be carefully considered. You must clearly explain why you do or don't agree with the feedback of your colleagues." (Q-P20)*

## Discussion

This study aimed to explore the critical features of a PA program that was shown to be effective in a previously published randomized controlled trial. The results clearly show that participants committed the most to learning tasks related to performance in the therapist role: performing the task, receiving peer feedback, external coach

feedback, and simulated patient feedback. Participants varied widely in the perceived learning value of subtasks related to performing the PT role, but agreed on the superior value of receiving peer feedback. In the next section, we will elaborate on these results. These results point to the importance of exposing observable behavior (PA) rather than expressing intended behavior (Case Discussion). Although exposure was associated with feelings of discomfort and performance stress, its impact on awareness of professional development was not questioned. This raises the question of how feelings of discomfort and stress can affect learning and change in professional practice. In the PT role, participants needed to make the transfer from implicit reasoning to explicit reasoning and from intentional behavior to observable behavior to allow for assessment and feedback. Bandura's social cognitive theory emphasizes that exposure is conditional to the development of mastery experiences, and mastery experiences are the most important source of information for the development of self-efficacy beliefs.

In turn, self-efficacy beliefs contribute significantly to performance improvement and motivation to change.[33] This notion is supported by the theory of planned behavior.[34] Bandura also points to the importance of the peer group in strengthening self-confidence through "vicarious" experiences provided by social models. The impact of modeling on perceived self-efficacy is strongly influenced by perceived similarity to the models (peers) and is considered to be more powerful than performance feedback.[35] Increased self-confidence might have helped participants to approach difficult tasks as challenges to be mastered rather than as threats to be avoided.

The foregoing explains how PA participants succeeded in raising self-efficacy beliefs despite feelings of performance stress, but does not explain why they showed superior test results on clinical vignettes in the trial (Table 1). High arousal levels are generally considered to have a negative impact on the quality of performance according to the Yerkes-Dodson law,[36] and PA participants' experiences supported that, as they contended that they had underperformed in the PA context. However, they must have processed the information in a way that enhanced retrieval and transfer of knowledge to the context of clinical vignettes. Studies addressing the influence of emotion on cognitive processing provide an explanation for this apparent contradiction. McConnel & Eva[37] conducted a literature review on the impact of emotion on the transfer of clinical knowledge and skills.

They conceptualized emotion by two dimensions: valence and arousal. Valence refers to the emotional state (e.g. positive or negative). Arousal refers to the level of activation. One of the findings was that emotional experiences are more likely to be mulled over than non-emotional experiences. This unintentional retrieval of emotional events might have strengthened memory traces of PA participants and facilitated the transfer to new clinical problems. Another view is presented by regulatory focus

theory,[38] which contends that receptiveness to feedback depends on "emotional arousal" rather than "emotional valence". Summarizing these considerations, the critical feature of PA might be attributed to the emotional involvement (either negative or positive) with performing the PT role. As feelings of failure do not contribute to the development of self-efficacy beliefs,[33] successful PA implementation should allow for coping with performance stress within or between the sessions. Training in the PT role and a safe learning environment might be crucial to enable the coping process.

Performance in the assessor role was perceived as a less powerful learning experience. However, it should be noted that the assessor role and the PT role cannot be considered as independent. Observing peer performance allowed observers to model the observed behavior, which might have contributed to reducing performance stress and triggering performance improvement. On a more unconscious level, participants might have profited from the activity of the mirror neuron system[39] that is capable of shaping the observed behavior to a virtual image of their intended behavior. In appraising their peers' performance, assessors needed to reason aloud, compare personal views with group views, and discuss performance standards. This may have provided peer assessors with the missing data for informed self-assessment.[20]



Regarding the role of the external coach in providing feedback, participants ranked peer feedback higher than coach feedback although coach feedback was valued because of its objectivity, its conciseness, and its receptiveness.

A comparable study on peer assessment in undergraduate PT education, in which students were asked to rank similar learning tasks, showed that students preferred teacher feedback to peer feedback.[17]

Professionals did not question the quality of peer feedback compared to coach feedback, but emphasized the importance of peers being involved in their professional development process. This finding is supported by situated learning theory,[40,41] which contends that the transfer of knowledge is hampered by the gap between the learning context and application context. Delivering the implementation program within communities of practice allows for co-constructing and tailoring knowledge to the personal learning needs.[41] In this respect, the coach remained an outsider.

Although the PA program was successful regarding its aim, the adoption of the program for knowledge transfer purposes should be carefully considered.

Firstly, some participants argued that the role-play format did not adequately reflect their authentic professional behaviors. This view is understandable, but compared to passive guideline dissemination, role-play aims to facilitate the transfer of scientific evidence to clinical practice, which it did, according to participant reports. As regards the use of peer role-play (low fidelity simulation) compared to standardized patients (high fidelity simulation), research in undergraduate education shows that

both tools provide a psychological safe area of practice, where mistakes are not critical.[42] Studies on student perceptions show that standardized patients are perceived as more effective than peers.[43,44] However, research evidence on learning outcomes remains inconclusive.[44,45] Compared to direct observation (work-place based assessment), the role-play format allows for standardizing the content of interest, creating an adequate case mix, and describing the key-features of health problems relevant to the guidelines.[46] Considering constraints in time and costs, peer role-play is the most feasible method. This conclusion is supported by a systematic review undertaken by Overheem et al.,[47] who evaluated the feasibility and effectiveness of six methods to assess physician performance. Secondly, some participants perceived the tight time schedule as stressing and preventing in-depth elaboration of the clinical problems. The PA program was designed to enhance the transfer from the learning context to the application context, as the transfer from one problem to another problem.[48] Yet, in an attempt to solve all the presented problems within time limits, the approach to learning might have been too superficial. Thirdly, performance in the PT role was perceived as challenging and sometimes even threatening. When conditions of psychological safety are not met, the effectiveness of PA might be questioned.[14]

## Strengths and limitations

This study provided rich data and convincing results. Because we clearly described the program design, its underlying theoretical constructs, and the critical features of successful guideline implementation, future program designers may profit from our results.

It can be argued that a limitation of the PA approach is the role-play of peers simulating patients. Although the choice of peers instead of standardized patients was defensible as argued above, and although the results show that their feedback was valued, additional training in the patient role might have increased the fidelity of the peers' performance.

Another limitation concerns the questionnaire and the interview guide. Questionnaire comments were reduced by the three tasks with the highest-ranking results. We compensated for this limitation by interviewing participants with contrasting ranking results. Nevertheless, because we did not focus on less instructive tasks in our interviews, we might have lost information that would have underpinned our results.

Finally, the generalizability of our results might be limited because all participants in this study were Dutch. Research shows that effective peer assessment practices are culture dependent.[23,24]

## Conclusions

The effectiveness of PA can be attributed to the structured and performance-based design of the program. Participants showed a strong cognitive and emotional commitment to performing the tasks related to the physical therapist role. That might have contributed to an increased awareness of strengths and weaknesses, and a motivation to change routine practice in the management of patients with low back pain.

Conditional to successful implementation is an environment omment where mistakes can easily be made, but in which the self-confidence of participants remains undamaged. Adjustment of the tight time schedule and the number of cases, providing more time to elaborate on problems and to recuperate from experiences, might improve the PA task design. However, attempts to improve the effectiveness of PA should not be limited to the modification of the PA tool. We recommend a shift in the feedback culture of PTs in primary care, from avoiding performance feedback to actively seeking feedback.

Future research should address the feasibility of PA as a tool to enhance bottom-up quality improvement and accountability to external stakeholders of PT care.

# References


1. Grol RP, Wensing M, Eccles MP, Davis DA: Improving Patient Care: The Implementation of Change in Health Care. 2nd edition. Chichester, West Sussex: Wiley-Blackwell; 2013.
2. Van der Wees PJ, Moore AP, Powers CM, Stewart A, Nijhuis-van der Sanden MWG, de Bie RA: Development of clinical guidelines in physical therapy: perspective for international collaboration. Phys Ther 2011, 91:1551–63.
3. Bekkering GE, Tulder MW Van, Hendriks EJ, Koopmanschap MA et al. Implementation of clinical guidelines on physical therapy for patients with low back pain: randomized trial comparing patient outcomes after a standard and active implementation strategy. Phys Ther 2005, 85:544–555.
4. Bekkering GE, Hendriks EJ, van Tulder MW, Knol DL et al. Effect on the process of care of an active strategy to implement clinical guidelines on physiotherapy for low back pain: a cluster randomised controlled trial. Qual Saf Health Care 2005, 14:107–112.
5. Van der Wees PJ, Jamtvedt G, Rebbeck T, de Bie RA et al. Multifaceted strategies may increase implementation of physiotherapy clinical guidelines: a systematic review. Aust J Physiother 2008, 54:233–41.
6. Harting J, Ruttten GM, Rutten ST, Kremers SP: A qualitative application of the diffusion of innovations theory to examine determinants of guideline adhrence among physical therapists. Physical Therapy 2009, 89:221–232.
7. Van Bodegom-Vos L, Verhoef J, Dickmann M, Kleijn M et al.: A qualitative study of barriers to the implementation of a rheumatoid arthritis guideline among generalist and specialist physical therapists. Phys Ther 2012, 92:1292–305.
8. Dannapfel P, Peolsson A, Nilsen P: What supports physiotherapists' use of research in clinical practice? A qualitative study in Sweden. Implement Sci 2013, 8:31.
9. Rutten GM, Kremers S, Rutten ST, Harting J: A theory-based cross-sectional survey demonstrated the important role of awareness in guideline implementation. J Clin Epidemiol 2009, 62:167–176.
10. Epstein RM: Self Monitoring in Clinical Practice. J Contin Educ Health Prof 2008.
11. Davis DA, Mazmanian PE, Fordis M, Harrison R Van, Thorpe KE, Perrier L: Accuracy of physician self-assessment compared with observed measures of competence. A systematic review. JAMA 2006, 296:1094–1102.
12. Eva KW, Regehr G: " I ' ll mever play professional football " and other fallacies of self-assessment. J Contin Educ Health Prof 2008, 28:14–19.
13. Sargeant J, Eva KW, Armson H, Chesluk B et al.: Features of assessment learners use to make informed self-assessments of clinical performance. Med Educ 2011, 45:636–647.
14. Eva KW, Armson H, Holmboe E, Lockyer J et al.: Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. Adv Health Sci Educ Theory Pract 2012, 17:15–26.
15. Mann K, van der Vleuten CP, Eva KW, Armson H et al.: Tensions in informed self-assessment: how the desire for feedback and reticence to collect and use it can conflict. Acad Med 2011, 86:1120–1127.
16. Baker R, Camosso-Stefinovic J, Gillies C, Shaw E, Chetaer F, Flottorp S, Robertson N: Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2015 Apr 29;4:CD00547.
17. Maas MJM, Sluijsmans DM, van der Wees PJ, Heerkens YF et al.: Why peer assessment helps to improve clinical performance in undergraduate physical therapy education: a mixed methods design. BMC Med Educ 2014, 14:117.
18. Van Dulmen SA, Maas MJ, Staal B, Kiers HJ, Bie RA, Van der Wees PJ, Nijhuis-van der Sanden MW: Effectiveness of peer-assessment for implementing a Dutch physical therapy low back pain guideline: a cluster randomized controlled trial. Phys Ther 2014, 94:1396–1409.
19. Mann K, Gordon J, MacLeod A: Reflection and reflective practice in health professions education: a systematic review. Adv Health Sci Educ Theory Pract 2009, 14:595–621.
20. Epstein RM, Siegel DJ, Silberman J: Self-monitoring in clinical practice: a challenge for medical educators. J Contin Educ Health Prof 2008, 28:5–13.

21. Pronovost PJ, Hudson DW: Improving healthcare quality through organisational peer-to-peer assessment: lessons from the nuclear power industry. BMJ Qual Saf 2012, 21:872–5.
22. Ramsey PG, Wenrich MD, Carline JD, Inui TS et al.: Use of peer ratings to evaluate physician performance. JAMA 1993, 269:1655–1660.
23. Lin SSJ, Liu EZF, Yuan SM: Web-based peer assessment : feedback for students with various thinking-styles. J Comput Assist Learn 2001, 17:420–432.
24. Topping KJ: Methodological quandaries in studying process and outcomes in peer assessment. Learn Instr 2010, 20:339–343.
25. Brehaut JC, Eva KW: Building theories of knowledge translation interventions: use the entire menu of constructs. Implement Sci 2012, 7:114.
26. Colquhoun HL, Brehaut JC, Sales A, Ivers N, Grimshaw J, Michie S, Carroll K, Chalifoux M, Eva KW: A systematic review of the use of theory in randomized controlled trials of audit and feedback. Implement Sci 2013, 8:66.
27. Ivers NM, Sales A, Colquhoun H, Michie S, Foy R, Francis JJ, Grimshaw JM: No more "business as usual" with audit and feedback interventions: towards an agenda for a reinvigorated intervention. Implement Sci 2014, 9:14.
28. Staal JB, Hendriks EJM, Heijmans M, et al. Richtlijn Lage-Rugpijn voor Fysiotherapie en Manuele Therapie [Guideline Low Back Pain for Physical Therapy and Manual Therapy]. Amersfoort, the Netherlands: Royal Dutch Society for Physical Therapy; 2013.
29. Janssen-Noordman AMB, Merriënboer JJG, van der Vleuten CPM, Scherpbier AJJA: Design of integrated practice for learning professional competences. Med Teach 2006, 28:447–452.
30. King N, Cassel C, Symon G: Using templates in the thematic analysis of texts. In Essent Guid to Qual methods Organ Res. 1st edition. London: Sage; 2004:256–270.
31. Huberman AM, Miles MB, Denzin NK, Lincoln YS: Data management and analysis methods. In Handb Qual Res. Sage Publications; 1994:428–444.
32. Qualitative research review guidelines - RATS [http://www.biomedcentral.com/authors/rats]
33. Bandura A, Locke E a.: Negative self-efficacy and goal effects revisited. J Appl Psychol 2003, 88:87–99.
34. Ajzen I: Nature and operation of attitudes. Annu Rev Psychol 2001, 52:27–58.
35. Usher EL, Pajares F: Sources of Self-Efficacy in School: Critical Review of the Literature and Future Directions. Rev Educ Res 2008, 78:751–796.
36. Teigen K: Yerkes-Dodson: a law for all seasons. Theory Psychol 1994, 4:525–547.
37. McConnell MM, Eva KW: The role of emotion in the learning and transfer of clinical skills and knowledge. Acad Med 2012, 87:1316–22.
38. Higgins TE: Beyond pleasure and pain. Am Psychol 1997, 52:1280–1300.
39. Iacoboni M: Mirroring People: The New Science of How We Connect with Others. 2nd edition. New York: Macmillan; 2009.
40. Lave J, Wenger E: Communities of practice. 1998.
41. Li LC, Grimshaw JM, Nielsen C, Judd M, Coyte PC, Graham ID: Evolution of Wenger's concept of community of practice. Implement Sci 2009, 4:11.
42. McCaghie WC, Issenberg BS, Petrusa ER, Scalese RJ: A critical review of simulation-based medical education research: 2003-2009. Med Educ 2010, 44:50–63.
43. Munshi F, Lababidi H, Alyousef S: Low- versus high-fidelity simulations in teaching and assessing clinical skills. J Taibah Univ Med Sci 2015, 10:12–15.
44. Bosse HM, Nickel M, Huwendiek S, Jünger J et al.: Peer role-play and standardised patients in communication training: a comparative study on the student perspective on acceptability, realism, and perceived effect. BMC Med Educ 2010, 10:27.
45. Mounsey AL, Bovbjerg V, White L, Gazewoord J: Do students develop better motivational interviewing skills through role-play with standardised patients or with student colleagues? Med Educ 2006, 40:775–780.
46. Farmer EA, Page G: A practical guide to assessing clinical decision-making skills using the key features approach. Med Educ 2005, 39:1188–1194.

47. Overheem K, Faber MJ, Onyebuchi AA, Elwyn G et al.: Doctor performance assessment development in daily practise: does it help doctors or not? A systematic review. and impact of a new system. Med Educ 2007, 41:1039–1049.
48. Norman G, Bordage G, Page G, Keane D: How specific is case specificity? Med Educ 2006, 40:618–623.
49. Grant AM, Franklin J, Langford P: The self-reflection and insight scale: a new measure of private self-consciousness. Soc Behav Pers 2002, 30:821–836.
50. Schön D: The Reflective Practitioner: How Professionals Think in Action. San Francisco: Jossey-Bass Inc; 1983.
51. Greene J, Azevedo R: A theoretical review of Winne and Hadwin's model of self-regulated learning: new perspectives and directions. Rev Educ Res 2007, 77:334–372.
52. Li LC, Grimshaw JM, Nielsen C, Judd M et al.: Use of communities of practice in business and health care sectors: a systematic review. Implement Sci 2009, 4:27.
53. Prochaska JO, Redding CA, Evers KE: Health behavior and health education. In Heal Behav Heal Educ theory, Res Pract. 4th edition. Edited by Glanz K, Rimer B k, Viswanath K. Wiley & Sons; 2008:97–121.

# Appendix

## Online questionnaire

1. The PA program consisted of several parts. The overview below shows the distinct learning tasks and subtasks. Please rank the eleven subtasks as presented from high to low learning value (1 = most learning value, 11 = least learning value).

| Overview tasks and subtasks | | | | |
|---|---|---|---|---|
| **Tasks** | | | **Subtasks** | **Rank** |
| Prepare Task | Study manual | 1 | Study PA procedure and guidelines | |
| Perform Task | Perform in PT role | 2 | Perform clinical task individually | |
| | | 3 | Receive peer feedback | |
| | | 4 | Receive external coach feedback | |
| | | 5 | Receive simulated patient feedback | |
| | | 6 | Receive written feedback and scores | |
| | Perform in assessor role | 7 | Observe peer performance | |
| | | 8 | Provide oral feedback | |
| | | 9 | Provide written feedback and scores | |
| | Perform in patient role | 19 | Simulate patient problem | |
| Evaluate Task | | 11 | Design and discuss change plan | |

2. Please motivate your choice for the three most instructive learning tasks.

| **Rank** | **Comment*** |
|---|---|
| 1 | |
| 2 | |
| 3 | |
| | *Characters unlimited |

## Interview guide

1) **What did you expect of the Peer Assessment (PA) program?**
   - Did you have personal learning goals? If so, can you describe them?
   - To what extent this program has met your expectations? Please explain.
   - The PA program aimed to enhance clinical performance of physical therapists in primary care. What are the strengths and weaknesses of PA, and why?

2) **How did you perceive the PA program, and how did it affect your daily practice?**
   - How did you perceive the PA sessions?
   - Can you remember a particular event that impressed you? If so, please describe.
   - When you look back on the PA process, did it affect your professional practice? If so, can you explain what has changed?
   - Do you think the PA process affected the professional practice of your colleagues? If so, can you explain what has changed?

3) **Which elements of the PA program to have the strongest learning value?**
   The PA program consisted of several parts. The overview below shows the learning tasks and subtasks. In the questionnaire you were asked to rank eleven subtasks as presented in the overview, from high to low learning value. You indicated that you perceived task *X* to have the strongest learning value. Can you explain why? Can you proceed to do the same for task *Y* and *Z*?

**Overview of learning tasks and subtasks**

| **Tasks** | | | **Subtasks** |
|---|---|---|---|
| Prepare Task | Study manual | 1 | Study PA procedure and guidelines |
| Perform Task | Perform in PT role | 2 | Perform clinical task individually |
| | | 3 | Receive peer feedback |
| | | 4 | Receive external coach feedback |
| | | 5 | Receive simulated patient feedback |
| | | 6 | Receive written feedback and scores |
| | Perform in assessor role | 7 | Observe peer performance |
| | | 8 | Provide oral feedback |
| | | 9 | Provide written feedback and scores |
| | Perform in patient role | 19 | Simulate patientproblem |
| Evaluate Task | | 11 | Design and discuss change plan |

1

# 5

# Supporting a person-centred approach in clinical guidelines. A position paper of the Allied Health Community – Guidelines International Network (G-I-N)

**Published as**

van Dulmen SA, Lukersmith S, Muxlow J, Santa Mina E, Nijhuis-van der Sanden MW, van der Wees PJ on behalf of the G-I-N Allied Health Steering Group. Supporting a person-centred approach in clinical guidelines. A position paper of the Allied Health Community – Guidelines International Network (G-I-N). Health Expectations 2015; 18(5): 1543-1558.

## Abstract

**Background** A person-centred approach in the context of health services delivery implies a biopsychosocial model focusing on all factors that influence the person's health and functioning. Those wishing to monitor change should consider this perspective when they develop and use guidelines to stimulate active consideration of the person's needs, preferences and participation in goal setting, intervention selection and the use of appropriate outcome measures.

**Objective** To develop a position paper that promotes a person-centred approach in guideline development and implementation.

**Design, setting and participants** We used three narrative discussion formats to collect data for achieving consensus: a nominal group technique for the Allied Health Steering Group, an Internet discussion board and a workshop at the annual G-I-N conference. We analyzed the data for relevant themes to draft recommendations.

**Results** We built the position paper on the values of the biopsychosocial model. Four key themes for enhancing a person-centred approach in clinical guidelines emerged: (i) use a joint definition of health-related quality of life as an essential component of intervention goals, (ii) incorporate the International Classification of Functioning, Disability and Health (ICF) as a framework for considering all domains related to health, (iii) adopt a shared decision-making method, and (iv) incorporate patient-reported health outcome measures. The position statement includes 14 recommendations for guideline developers, implementers and users.

**Conclusion** This position paper describes essential elements for incorporating a person-centred approach in clinical guidelines. The consensus process provided information about barriers and facilitators that might help us develop strategies for implementing person-centred care.

## Background

Health service providers in clinical practice are encouraged to adhere to evidence-based guidelines, which are considered important tools for quality improvement and health outcomes.[1] Evidence-based medicine is based on the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients,[2] whereas evidence-based practice is the translation of the evidence into health services. Best-practice health services also need to take into account the contextual factors that affect the person's health condition. Contextual factors include internal personal factors and external environmental factors. Personal factors derive from the individual and include gender, age and education. Environmental factors make up the physical, social and attitudinal environment of a person; they include family support and the health service system.[3] Contextual factors may affect a person's functioning and can be either a barrier to or a facilitator for the improvement of personal health outcomes within a health services system.

Historically, health-care decision making has been based on recommendations from quantitative medical research and knowledge that focus on medical diagnosis, impairments and treatment. However, the understanding of health conditions and a person's functioning, disability and health has changed from a medical perspective focusing on the individual's physical aspects to a biopsychosocial perspective that recognizes the relationship between the individual and other related context.[4] The biopsychosocial model posits that biological, psychological and environmental or social factors all influence the individual's functioning and health outcomes.[5] From the biopsychosocial perspective, health-care professionals should consider not only disease processes and the biomedical aspects of the person, but also the environmental and personal factors, the person's quality of life (QoL) and participation in all major life areas including making decisions and the choice of and control over his/her health and the use of health services. This implies a paradigm shift in the awareness of the roles of health professionals and persons with a health problem. Health professionals should understand their role as a health service provider, and persons with (or at risk for) health problems are service users. Needs and expectations differ between persons, and it takes time to change the paradigm from a paternalistic approach and perspective of a 'patient' who can be viewed as a passive recipient of the health service to a user perspective where there is choice and control. In this manuscript, we will use terms to reinforce this paradigm shift and use health rather than health care, health service provider rather than health-care professional and service user rather than patient, except in circumstances where these words are used in definitions, quotes or specific references.

## Person-centred approach

The terms 'person-centred' and 'patient-centred' are often used interchangeably in health services. In this paper, we have adopted the term 'person-centred' to better reflect the biopsychosocial model. In a person-centred approach, a person (the service user) is valued as an active participant of the health service.

Evidence-based interventions should be adapted to meet individual needs and preferences where possible. This approach provides insight into factors related to a disease and facilitates interventions congruent with personal values, beliefs and environmental conditions of the health service user. The adoption of a person-centred approach in a health services system can influence personal factors such as patient satisfaction, motivation, adherence to therapy and thus health outcomes.[6–9] Elements of person-centred care include communication and relevant information, as well as shared decision making (SDM) and self-management support.[10] The Institute of Medicine includes patient-centred care as one of the six essential aims for improving health care.[11] Although person-centred care is advocated in clinical practice, its implementation is constrained due to variation in definitions, perceived barriers to valid and reliable outcome measurements and the inherent difficulty with shifting traditional patterns of interaction.[10,12–15]

## Clinical practice guidelines

Clinical practice guidelines are important tools for synthesizing evidence and translating research findings into practice. Their purpose is to assist health service providers and service users in health behaviour and decision making. Clinical practice guidelines should specifically aim at integrating the complex interplay of science with the provider's experience, reasoning and judgment, and the person's values and preferences in conjunction with consideration of both the context of the health service and the person's environment.[16] Guidelines potentially support a person-centred approach by guiding the health service provider in considering all relevant domains of the person's health, facilitating individualized and meaningful goal setting, recommending appropriate intervention strategies and using outcome measures that monitor change in the person's health and functioning.

## Guidelines International Network

The Guidelines International Network (G-I-N) is a global network that supports evidence-based health care and improvement of health outcomes by promoting the development, implementation and use of clinical practice guidelines internationally. Two of the aims of the Allied Health Community of the G-I-N are to promote person-centred health services and to promote health-related QoL activities –objectives for prevention, assessment, intervention and evaluation – in multidisciplinary clinical guidelines. The aims of our study are to identify some of the key themes necessary

to make guidelines more person-centred and to develop a position paper to promote these elements for a person-centred approach in guideline development and implementation.

# Method

## Study design

We adopted a consensus method, using a mix of techniques to draft the position statement. The method consisted of three stages for identifying key themes linked to person-centredness for the position statement: (i) a nominal group technique (NGT) used in the G-I-N Allied Health Steering Group (three rounds) combined with a focused literature search, (ii) an Internet discussion board for G-I-N members, and (iii) a workshop at the G-I-N 2011 annual conference.

The NGT involves a structured approach with discussion between the participants and prioritization of the themes.[17,18] The G-I-N Allied Health Steering Group participated in three nominal group rounds in teleconference, which lasted 1.5 h each. These sessions took place between December 2010 and June 2011. The G-I-N Allied Health Steering Group consisted of nine members from four countries. Table 1 provides an overview of the members' characteristics.



We used similar approaches for each of the three rounds. The modified NGT consisted of the following stages: introduction and explanation, silent generation of ideas, sharing of ideas, group discussion and ranking of ideas. The choice of the NGT directly involved the participants in both data collection and analysis, which helped ensure that their findings and interpretations accurately reflected their thoughts. A topic guide was used for semi-structured discussions in the stages of sharing ideas and group discussions. In between the nominal group sessions, we explored the literature about the emerging themes to develop clarity and definitions and to see how we could fit these themes into clinical practice guidelines. We searched the PubMed, Cinahl, and EMBASE databases using combinations of the following keywords: biopsychosocial model; International Classification of Functioning, Disability and Health (ICF); shared decision making; health related QoL; outcome measures; clinical practice guidelines.

Before each succeeding session, PW and SD provided an overview of the results of the previous session together with a summary of the findings from the literature review. We used the refined themes as the basis for further questioning, discussion and consensus development with the broader G-I-N community. Given the international character of person-centredness research, we initiated an Internet group discussion with pre-defined questions about the identified themes. The

**Table 1** Characteristics of Allied Health Group members and participants of the internet discussion board

| Variable | Members of the Allied Health Steering Group | Respondents on the internet discussion board |
|---|---|---|
| Number of participants | 9 | 14 |
| Sex male/female | 1/8 | 7/7 |
| **Healthcare background** | | |
| Physical therapist | 5 | 1 |
| Nurse | 2 | 1 |
| Occupational therapist | 1 | |
| Health scientist | | 1 |
| Human movement scientist | 1 | |
| Psychiatrist | | 1 |
| Pharmacist | | 1 |
| General internist | | 1 |
| General practitioners | | 2 |
| Librarian | | 1 |
| Unknown | | 5 |
| **Country** | | |
| Netherlands | 4 | 5 |
| United Kingdom | 2 | 4 |
| Germany | | 3 |
| Belgium | | 1 |
| Canada | 2 | |
| Australia | 1 | 1 |

proposed methodology facilitated joining the discussions in a rapid and convenient manner. The targeted participants were the 180 G-I-N members. We used a generic approach in inviting all G-I-N members by email to join the Internet discussion board. We asked the participants to respond whether they had specific knowledge about one of the themes. In addition, we invited 10 experts external to the G-I-N who were knowledgeable about the themes to contribute to the Internet discussion. These experts were researchers and health service providers with specific expertise in the ICF, QoL research, SDM methods or patient-reported outcome measures. All participants were asked to state whether they endorsed the themes and to respond to the related questions. The participants were encouraged to add comments and

suggestions in contributing to the discussion about any of the themes. The website was open for 2 months (July and August 2011).

We organized a workshop at the annual G-I-N conference in Seoul in August 2011. The 250 participants of the international conference could choose from seven parallel programmes at the proposed time of the workshop. The aim of the workshop was to reflect on the themes and to initiate further discussion. During the workshop, an overview of the aim and method was provided, followed by the information about the themes, definitions and the questions.

The participants were asked to respond to questions identical to those for the online discussion board to encourage further discussion and consensus. The discussions at the workshop were audio-recorded and transcribed.

The G-I-N Allied Health Steering Group used information from all stages of the development to finalize the position statement by providing recommendations for guideline developers, guideline implementers and health service providers. Figure 1 presents an overview of the study design.

### Analysis

After the website closed, all comments from the discussion board were merged with the transcripts of the discussions from the workshop. We analysed the data to identify individual concepts of health professionals' perspectives. Similar arguments from different respondents were grouped, but individual remarks and disagreements were pointed out as well. One researcher (SD) selected key arguments and essential elements within each expert's reactions, and a second researcher (PW) validated all identified elements. The key arguments or topics raised for each theme were used to structure the results and to develop draft recommendations. The Allied Health Steering Group discussed and amended the draft recommendations.



## Results

### Nominal group sessions and the targeted literature search

The first nominal group session aimed at defining the scope of the position paper and topics for inclusion. This resulted in the biopsychosocial model for developing the position paper. The participants then shared ideas about how to integrate the biopsychosocial model into clinical guidelines and about which themes were relevant for incorporation in the position paper. The group of participants agreed on a shortlist of selected themes for further exploration in preparation for the second session. The selected themes were refined and grouped within themes during the second session. At the end of this session, we identified four themes to enhance a person-centred approach in guidelines: (i) health-related quality of life (HRQoL), (ii) ICF framework,

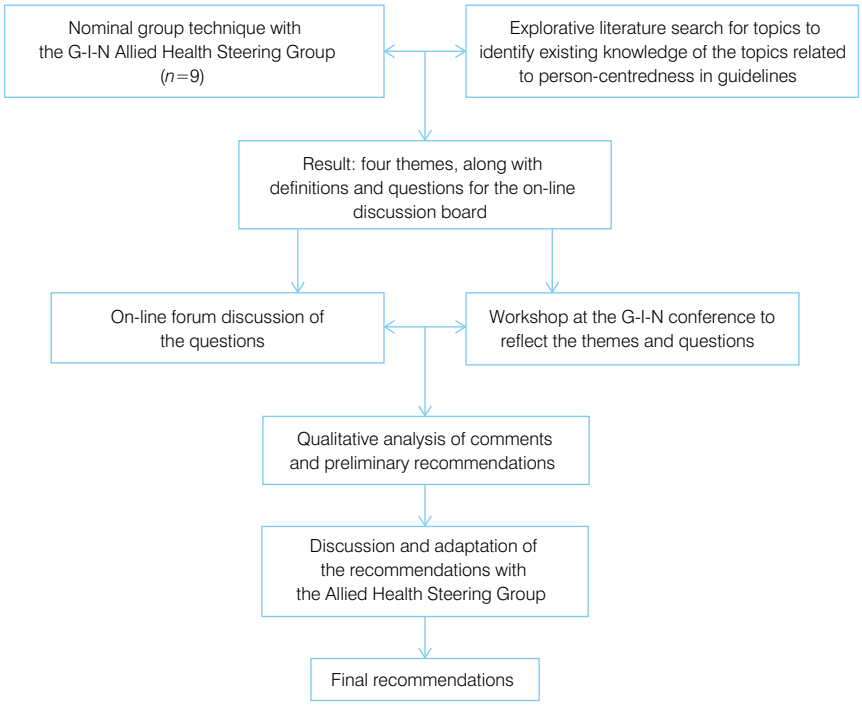


**Figure 1** Overview of the study design

(iii) SDM method, and (iv) HRQoL outcome measurement. In the third session, we discussed the selected themes and definitions and considered whether they are essential to enhancing the person-centred approach in guideline development and implementation. This led to the selection of four key themes for further discussion via the Internet discussion board. The participants then shared ideas about the content of the Internet discussion board, for example what information should be included in addition to the definitions, what questions should be asked and who should be invited. This resulted in a structure for the Internet discussion board: provide a clear definition of each theme with background information and ask questions to initiate discussion for each of the key themes. Box 1 shows the definitions and questions for the discussion board.

## Key themes

The Allied Health Steering Group identified four themes that we considered essential to enhancing a person-centred approach in guideline development and implementation. The themes were as follows:

**Box 1** Themes and questions on the internet discussion board

| Theme | Definitions and questions |
|---|---|
| HRQoL | Definition: the person's perception of the impact of his disease and its intervention(s) in his daily life – physical, psychological, and social functioning and well-being.<br><br>Question:<br>- Should the concept of HRQoL be used in all guidelines? |
| ICF | Definition: the ICF as a good classification and framework for functioning, disability, and health-related domains. These domains are classified by body, functions, and structures; activities and participation; and personal and external perspectives.<br><br>Questions:<br>- Do guideline developers, guideline implementers, and service providers generally accept the ICF framework?<br>- Does the ICF provide a good framework for promoting quality of life in guidelines? |
| SDM | Definition: decision-making jointly shared by patients and their health service providers is a move toward seeing patients as having a central role with the aim of strengthening and empowering them to express their values and preferences, to ask questions, and to participate actively in their own health.<br><br>Questions:<br>- Do guideline developers, guideline implementers, and health service providers generally accept the SDM method?<br>- Should decision aids be recommended in guidelines? |
| Quality-of-life outcome measurement | Definition: Quality-of-life measures capture patients' perspectives of their diseases and interventions, their perceived need for health services, and their preferences for interventions and outcomes. Such measures are useful for goal setting and decision-making for the individual patient.<br><br>Questions:<br>- Do patient-reported outcome measures contribute to promoting quality of life in guidelines?<br>- Should guideline developers aim at including measures related to quality of life in all guidelines? |

HRQoL: Health-related quality of life; ICF: International Classification of Functioning, Disability and Health; SDM: Shared Decision-making

### *Use a single definition of QoL as an essential component of guideline recommendations*

Although QoL and HRQoL are often used interchangeably, they are in fact different constructs.[19] 'Quality of life' has been defined as the capacity of an individual to achieve his/her life plans, or as the difference, at a particular point in time, between the hopes and expectations of an individual and his/her present situation.[20] A generally accepted definition of HRQoL is the person's (patient's) subjective perception of the impact of his/her disease and its intervention(s) on his/her daily life – physical, psychological and social functioning and well-being.[21] In this paper, we use HRQoL because, by definition, it reflects the biopsychosocial model and a person-centred care approach, and it seems to be the most appropriate in the context of health services.

### *Use a construct that covers all domains related to the person's health*

The World Health Organization (WHO) has developed the ICF to provide unified, international and standardized language to describe and classify health and health-related domains. This furnishes a common framework for health and health-related outcome measurement.[3,22] The ICF incorporates three main domains of health and functioning: body functions and structures, activities and participation, and contextual factors (environmental and personal factors) and is aetiologically neutral. The ICF provides a framework for understanding health, disability and functioning, as well as a classification system for use in clinical practice and research.[23–27] Figure 2 provides an overview of the domains of the ICF. The ICF creates a foundation for considering all aspects of the person's health condition during clinical assessment and for incorporation in guidelines; it is useful for diagnosis, goal setting, intervention selection and evaluation because it takes into account the contextual factors.[3,24,28–32]

### *Promote the use of a SDM method*

Shared decision making promotes informed choices, ensuring that person–professional interaction is responsive to individual needs, values and priorities.[33] It is a process in which the health service provider and the person with health problems as service user make a choice jointly;[34] SDM is pivotal to person-centred care.[35] User involvement at each stage in the development of a clinical guideline is essential to facilitate the adoption of SDM methods within a guideline. Involving service users and representatives in developing a guideline will ensure that the views, experiences and interests of users are considered in the guideline. The guideline needs to promote SDM methods at the level of the health service consultation. In clinical practice, SDM promotes active participation in intervention decision making, self-efficacy and self-management.

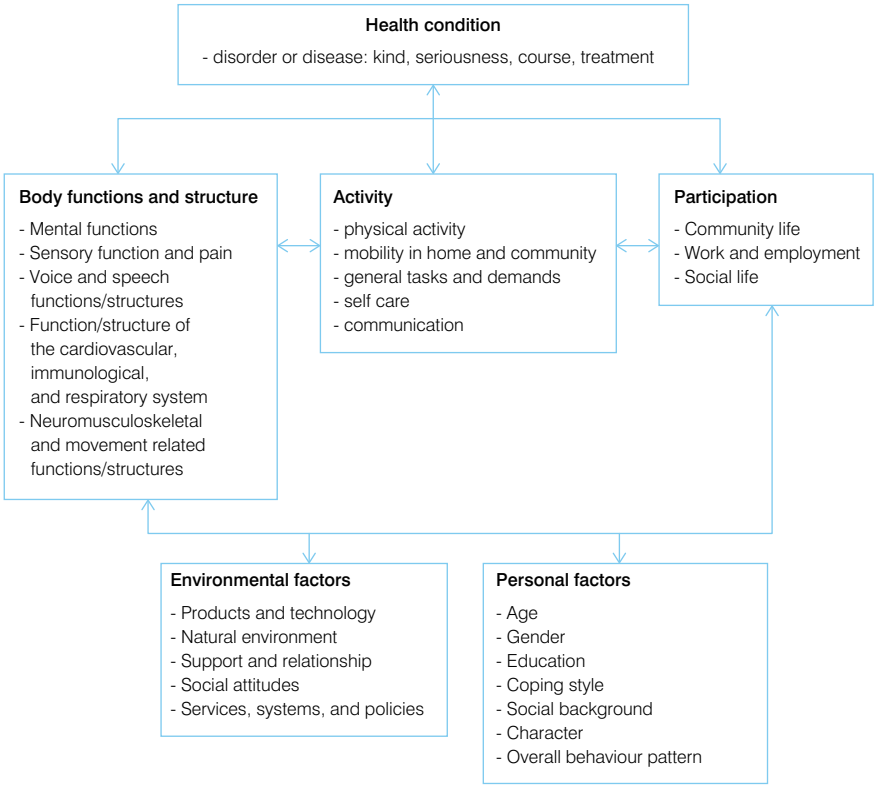


**Figure 2** Domains of the ICF

### *Monitor health outcomes*

For comprehensive assessment of the benefits of an intervention, it is essential to provide evidence of the impact of the health services on the person's health condition in all domains of health. Guidelines should therefore promote monitoring of health outcomes in clinical practice.

The HRQoL measurement refers to aspects of the individual's physical, emotional and social well-being,[36] and it reflects all domains of health articulated in the biopsychosocial model. Most HRQoL measurements are based on patient-reported outcome measures (PROMS), which are used in clinical practice to gain more insight into the way a person perceives his/her health and the impact of interventions on his/her QoL.[37] Frequently used instruments of HRQoL measurement are generic (e.g. the Short Form (36) Health Survey) or disease specific (e.g. the Asthma Quality of Life Questionnaire). The benefits of measuring PROMS in clinical practice include

facilitating communication about issues that are important to the patient as service user (which promotes SDM), evaluating the response to the intervention and enabling continuous assessment of quality of care.[38,39]

## Internet discussion board and workshop

We received 38 comments from 17 participants; 14 participants on the Internet discussion board and three participants in the workshop. There were comments on each theme within the proposed framework.

### *General comments*

All participants confirmed the importance of a person-centred approach, and some participants acknowledged that a person-centred approach is currently not routinely integrated into guidelines and clinical practice.

> *"The change from a medical perspective towards a more biopsychosocial model has not been achieved yet' and 'Patient perspectives are needed in study designs, guideline development and implementation in clinical practice."*

### *Health-related quality of life*

Most participants embraced the idea of including the concept of HRQoL in guidelines.

> *"If we purport to use evidence-based clinical practice and SDM, we must incorporate HRQoL as a central feature of guideline development. Engaging patients at the initial inception of the guideline and then at all following stages is critical."*

Some participants suggested that improved QoL should be addressed as an ultimate goal in guidelines focusing on chronic diseases and health conditions with a high HRQoL impact.

> *"The concept of HRQoL should be used in most guidelines, certainly for guidelines about chronic conditions and sometimes for non-chronic conditions or problems too."*

> *"It is probably not relevant to include HRQoL in guidelines dealing with acute conditions when it can have but little long-term effect."*

Several participants emphasized the importance of HRQoL in guideline development and clinical practice in relation to goal setting for the intervention and the person's own responsibility for his/her health, whilst cultural and cost aspects should be taken into account.

*"The main goal is quality of life, but in clinical reasoning, more goals can be formulated to reach this main goal… this means that QoL measurement should be related to the common goal of both the patient and the provider."*

*"Physicians themselves must adopt the patient's HRQoL as one of their major treatment aims and discuss HRQoL issues with the same priority as objective patient data."*

*"When the concept of HRQoL is used in guideline development, one should also pay attention to the dimension of "culture", "the spiritual level", and "economic aspects", depending on the topic or focus."*

Other participants noted the need for further research for the usability and applicability of HRQoL measures, which is a prerequisite for incorporation in guidelines.

*"As guideline developers, we are also dependent on whether the concept of HRQoL is being used in the related healthcare and medical research – and of course whether it is being used by healthcare professionals themselves in daily practice. Building up evidence and developing guidelines is only possible on the basis of good research and implementation in daily practice."*



### *International Classification of Functioning, Disability and Health*

All respondents affirmed that the ICF framework provides guidance in clinical practice for considering all relevant dimensions of health, and it takes into account the facilitators for and barriers to recovery.

*"Personal and environmental factors are the most important ones in client-centred care and shared decision-making."*

The ICF framework and coding system are generally not used in guidelines, except for some examples in the field of rehabilitation, occupational therapy and physical therapy. A few participants noted the dominance of disease-related clinical outcomes and the complicated and time-consuming coding system as reasons for the limited use of the ICF, although the framework is appropriate for consideringall domains of health.

*"Most questions developed by guideline teams and researchers focus on the body functions and activity because these are deemed easier and more accurate measurement domains. This view is often echoed in clinical assessments and treatment plans. The ICF would be an excellent and appropriate framework, but it may require further "education" of peers to place it in its appropriate perspective."*

### *Shared decision making*

The participants perceived SDM positively as a way to incorporate personal values and intervention preferences because it will make persons with questions and needs active members in goal setting and empower them to take an interest and responsibility for their health.

> *"Healthcare providers increasingly recognize the relevance of patient involvement in decision-making and patient activation in care in general, especially when self-management is required."*

The participants noted increased recognition of the facilitating role guidelines might have in SDM by taking into account information about consequences of interventions, but also by considering whether decision aids and the development of patient versions (service users) were included in the guidelines.

> *"Guidelines should move to a situation where there are patient versions that act as if they are decision support tools."*

However, according to some participants, SDM is not widely implemented in clinical practice because health service providers lack skills and use their own perceptions and definitions about the involvement of patients as service users in medical decision making.

> *"Providers do not always have the skills to activate patients in decision-making."*

> *"Shared decision-making is not suitable for all patients. Some of them will say, "Doctor, please tell me what to do… You know what's the best option for me."*

However, other participants state that there are barriers to the understanding of the SDM method in the guideline develop team.

> *"I'm not sure that SDM has already been implemented in the minds of guideline developers', and 'If it were accepted, then the concept of declaring reasonable options would be common."*

Comparative information in meaningful and accessible format is a prerequisite for SDM, which is still a challenge.

> *"Comparative, unbiased information is often not available, not even to the healthcare professional. So how can we expect doctors to be able to inform their patients correctly?"*

*"If decision aids are produced by commercial organizations (e.g. software companies) you need to be particularly careful about bias."*

### *Outcome measures of health-related quality-of-life*

The participants highlighted the relevance of using HRQoL PROMs in addition to clinical measures to facilitate the dialogue between the person and professional about what is important to them and to the relevance and effectiveness of the intervention. Therefore, outcome measures that are specific and relevant to the targeted population must be incorporated into the guidelines.

*"Outcome measures can certainly contribute and should be included – but they should be linked to the key clinical questions and the key messages of the guideline."*

The participants emphasized the importance of using evidence-based instruments that are suitable for routine clinical practice in order to incorporate HRQoL outcome measurements into guidelines and to facilitate implementation in clinical practice. They also emphasized the principle that measurement results should be immediately available to the individual.

*"For many conditions, there is no good evidence whether one should use a disease-specific instrument or a generic one and whether one of the available instruments is preferable in the given circumstances."*

*"There needs to be an understanding of the tools and knowledge of how to interpret the result; for example, to know whether any change is meaningful."*

## Discussion

This position paper describes four themes that we perceived to be important for a person-centred approach in guidelines. In guideline development, person-centred care means promoting the active involvement of the person in taking an active role and responsibility for his/her own health, ensuring better consideration of the person's needs, preferences and context and evaluating relevant health outcomes. The ultimate aim of health services should be to increase or maintain the person's HRQoL. All domains within the ICF framework are potentially relevant and should be considered for incorporation in the guideline during development.

This framework is also important for developing an understanding of the interactions between all domains of the individual's health in clinical practice. The

use of the ICF and HRQoL measures will promote the use of SDM methods and provide opportunities to map scenarios for individualized person-centred goals and realistic intervention strategies.

The personal perspective is incorporated at various levels: by taking personal factors into account at the ICF level, by involving the service user at the SDM level and by considering the HRQoL PROMS for feedback from the individual about his/her perceived health and the effect of interventions.

### *The impact of a person-centred approach*

We argue that person-centred care focuses sharply on participation of the person in clinical decision making by taking into account his/her perspective and by tailoring health services to the needs and preferences of the person. This approach has important implications for the outcome measures and the evaluation of the effectiveness of person-centred interventions. The implication is that outcome measures should be related to the perceived effect on HRQoL.

Patient-reported outcomes may also have benefits for improving problem detection, defining intervention goals, user–provider communication, SDM and assessing the effectiveness of the intervention.[40–43] This is especially relevant for managing chronic diseases (where evidence-based health services require increasingly complicated and expensive interventions) and for the on-going support for self-management.[44] Guideline developers must be aware of these benefits and should search for PROMS that are relevant to the targeted population.

### *Challenges for a person-centred approach in clinical guidelines*

The translation of evidence into clinical practice guidelines has been widely adopted in modern health services, although adherence to guidelines is extremely variable.[45–48] Evidence-based medicine offers the best available evidence for the most effective intervention, mostly via randomized clinical trials. Clinical trials typically use strict inclusion criteria and outcome measures focused on specific diagnosis and the biomedical condition, and it frequently neglects relevant factors in decision making.[49] Diverse barriers hinder translation of the results of clinical trials into clinical practice. These barriers include individual characteristics, interventions and health service providers, as well as environmental and practical issues, which are embedded in everyday 'real life' feasibility issues that prevent strict adherence to guidelines.[50–52]

A person-centred focus has a holistic, biopsychosocial perspective: there is a sharp focus on participation in the diagnostic procedure, goal setting, intervention selection and prioritization of outcomes from both the provider and the user. Health service providers perceive tension between the need to respect personal preferences and the pressure to strictly apply guidelines. It is a challenge for health service providers to integrate both paradigms in clinical decision making,[2,49] and to resolve

what are still sometimes seen as competing issues. The participants pointed out the dominance of the medical model in curricula and research as a limitation of the biopsychosocial perspective. The relative unfamiliarity of the framework of the ICF in the results of this study and in the literature confirms this.[53] Providing more information about the ICF framework might be a good approach for making the shift from the medical or social model to a solid application of the biopsychosocial model in a SDM approach.

Guidelines have yet to integrate evidence and tools that could enable the health service providers to consider personal values and preferences and to discuss alternative interventions.[23] The use of HRQoL PROMS in clinical practice is limited.[37,54,55] Our study participants recognize the benefits of these measures because they add a unique value and can inform person–professional decision making. The HRQoL PROMS provide information about how alternative interventions compare in terms of their benefits and risks for goal setting and about selecting the best intervention.[42] Guideline developers should include outcome measures in such a way that they facilitate the translation of generalized, evidence-based recommendations into individualized preferencebased decisions for specific groups of people. This means that HRQoL PROMS should be included in guidelines, in addition to clinical measures of biomedical status (e.g. laboratory, radiographic and physical examination). The participants noted guideline developers' difficulties in choosing the best measure (e.g. should it be based on the best evidence or the goal of including the outcome measure?) and barriers in clinical practice such as time constraints, interpreting results and the link to goal setting. These and other barriers such as costs, applicability to individuals, outcome measurement skills and possible lack of motivation for collecting and using PROMS must be considered.[37,54,56–60] Training health professionals may be necessary to overcome such barriers, and this training should be part of the guideline implementation.

Involvement of service users is important not only in clinical practice, but also in guideline development. The AGREE Research Trust has developed the Appraisal of Guidelines for Research and Evaluation (AGREE II) instrument for guideline development, reporting and assessment. One of the quality criteria concerns patient involvement as service users in the development of the guidelines.[61,62] Recently, the G-I-N proposed a key set of components for guideline development that emphasizes the importance of including personal preferences of users in the rating of evidence and recommendations.[63] Our participants have reported that SDM is not widely adopted in guideline development and clinical practice. Active involvement of patients as service users in the early stages of developing the guideline is important for addressing themes that are important to these users[64] and leads to the identification of issues that may not have otherwise been considered.[65,66] Engaging service users is essential for recognizing the impact of the health condition or disease

on functioning and QoL and also to get insight into the possible options and active participation in the intervention strategies. Guideline recommendations at the decision-making level and decision aids for the service user can make clinical practice guidelines more sensitive to the person's preferences. Decision aids increase people's involvement and empowerment, improve knowledge and promote a realistic perception of outcomes. They also appear to have a positive effect on user–provider relationships.[67–69]

## Limitations

This study has several limitations. First, there was little response to the Internet discussion board and at the conference. Whilst we invited all G-I-N members, many may not have considered themselves specific experts on the themes, but it may also reflect the G-I-N membership. The focus of many G-I-N members may be the content of clinical guidelines rather than joining a meta-level discussion as we conducted in our study. Another reason for the low response could be the relatively short period (July and August) and the time of the year, which is the period when many G-I-N members in the northern hemisphere may have taken their vacation. The participants were mostly clinicians (e.g. physicians and allied health professionals), and their comments were on the individual level of interaction between person and clinician. The principles involved may be applicable to a person-centred approach in public health as well. Second, whilst we focused our literature search on key themes and definitions, it was not an exhaustive or systematic review. As a result, we may have missed relevant definitions and interpretations of key themes beyond the four identified themes for improving person-centred care in guidelines. Despite this, the GI-N Allied Health Steering Group considers it worthwhile to start the challenge to guideline developers, users and implementers to incorporate and consider these four most relevant themes identified in this study.

## Recommendations

The G-I-N's Allied Health Community seeks to promote best practice by adopting a person-centred approach in developing guidelines and their implementation. This requires an integrated approach that considers the complex interaction of the relation of the underlying disease with the functioning of the individual in his/her social context and the inclusion of individual values and preferences. Table 2 presents our final recommendations for guideline developers, guideline implementers and health service providers for enhancing a person-centred care approach throughout guideline development, as well as during guideline implementation. The next steps for further research should be to analyse current guidelines for evidence of person-centred care approaches and to explore variations in guideline recommendations regarding their integration.

**Table 2** Recommendations for guideline developers, guideline implementers, and health care service providers

| |
|---|
| **Guideline developers** |
| • Explicitly describe HRQoL in the guideline for exploring the needs and values of the patient as a service user, goal setting, and outcome measures.<br>• Use the ICF framework for the guideline to describe all relevant dimensions of the individual's functioning, health, and disability, and consider the interacting factors.<br>• Define outcome measures that address the intervention's effectiveness in the relevant ICF domains.<br>• Incorporate service users as active members in developing the guideline.<br>• Include decision aids in guidelines, together with the evidence-based information that underpins shared decision-making.<br>• Provide a service user version of the guideline or make patient information available in plain language, and include information about the advantages and disadvantages of interventions and possible choices. |
| **Guideline implementers** |
| • Present tailored information about the professional responsibility to share information with service users, to ask about their needs and values, to offer different intervention options if appropriate, and to engage them patients in intervention and outcome measures.<br>• Provide accurate support tools based on the ICF to describe health factors explicitly.<br>• Use adequate linking rules to connect health problems to intervention goals and outcome measures for HRQoL, related to the ICF domains. |
| **Healthcare service providers** |
| • Recognize the individual's HRQoL in diagnosis, goal setting, and intervention selection.<br>• Consider all relevant dimensions of the individual's health functioning and disability and all interacting factors in line with the ICF framework.<br>• Use service user reported health outcomes, based on quality-of-life measurements, for shared goal setting, intervention selection, and user participation and evaluation.<br>• Provide and share clear and evidence-based information with the individual and others who are involved in their intervention decisions.<br>• Tailor the information to the individual and include advantages and disadvantages of intervention options. |

HRQoL: health-related quality of life; ICF: International Classification of Functioning, Disability and Health

# References


1. Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001; 286(20): 2578-85.
2. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312(7023): 71-2.
3. World Health Organization W. Towards a common language for functioning, disability and health - ICF. Geneva: WHO, 2002.
4. Kraus de Camargo O. Systems of care: transition from the bio-psycho-social perspective of the International Classification of Functioning, Disability and Health. Child Care Health Dev 2011; 37(6): 792-9.
5. Engel GL. The need for a new medical model: a challenge for biomedicine. Science 1977; 196(4286): 129-36.
6. Henbest RJ, Stewart M. Patient-centredness in the consultation. 2: Does it really make a difference? Family Practice 1990; 7(1): 28-33.
7. Henbest RJ, Fehrsen GS. Patient-centredness: is it applicable outside the West? Its measurement and effect on outcomes. Family Practice 1992; 9(3): 311-7.
8. Kinnersley P, Stott N, Peters TJ, Harvey I. The patient-centredness of consultations and outcome in primary care. Br J Gen Pract 1999; 49(446): 711-6.
9. Stewart MA. Effective physician-patient communication and health outcomes: a review. CMAJ 1995; 152(9): 1423-33.
10. Robinson JH, Callister LC, Berry JA, Dearing KA. Patient-centered care and adherence: definitions and applications to improve outcomes. J Am Acad Nurse Pract 2008; 20(12): 600-7.
11. IOM. Crossing the Quality Chasm: A New Health System for the 21st Century. In, Washington, D.C: Institute of Medicine, 2001.
12. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. Health Res Policy Syst 2006; 422.
13. Krahn M, Naglie G. The next step in guideline development: incorporating patient preferences. JAMA 2008; 300(4): 436-8.
14. Mead N, Bower P. Patient-centredness: a conceptual framework and review of the empirical literature. Soc Sci Med 2000; 51(7): 1087-110.
15. Gillespie R, Florin D, Gillam S. How is patient-centred care understood by the clinical, managerial and lay stakeholders responsible for promoting this agenda? Health Expect 2004; 7(2): 142-8.
16. IOM. Clinical practice guidelines we can trust. In, Washington, D.C.: Institute of Medicine, 2011.
17. Potter M, Gordon S, Hamer P. The physiotherapy experience in private practice: the patients' perspective. Aust J Physiother 2003; 49(3): 195-202.
18. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995; 311(7001): 376-80.
19. Davis E, Waters E. Re: A structured review of studies on health-related quality of life and economic evaluation in pediatric acute lymphoblastic leukemia. J Natl Cancer Inst 2005; 97(23): 1786-7.
20. Bjornson KF, McLaughlin JF. The measurement of health-related quality of life (HRQL) in children with cerebral palsy. Eur J Neurol 2001; 8 Suppl 5183-93.
21. EMEA. Reflection Paper on the Regulatory Guidance for the Use of Health Related Quality of Life (HRQL) Measures in the Evaluation of Medicinal Products. In, edCHMP, London: UK: European Medicines Agency, 2005.
22. Cieza A, Stucki G. The International Classification of Functioning Disability and Health: its development process and content validity. Eur J Phys Rehabil Med 2008; 44(3): 303-13.
23. Boivin A, Legare F, Gagnon MP. Competing norms: Canadian rural family physicians' perceptions of clinical practice guidelines and shared decision-making. J Health Serv Res Policy 2008; 13(2): 79-84.
24. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: an update based on lessons learned. J Rehabil Med 2005; 37(4): 212-8.
25. Stucki G, Kostanjsek N, Ustun B, Cieza A. ICF-based classification and measurement of functioning. Eur J Phys Rehabil Med 2008; 44(3): 315-28.

26. Fayed N, Cieza A, Bickenbach J. Illustrating child-specific linking issues using the Child Health Questionnaire. Am J Phys Med Rehabil 2012; 91(13 Suppl 1): S189-98.
27. Fayed N, Cieza A, Bickenbach JE. Linking health and health-related information to the ICF: a systematic review of the literature from 2001 to 2008. Disabil Rehabil 2011; 33(21-22): 1941-51.
28. Heerkens Y, Bougie T, Claus E. The use of the ICF in the process of supplying assistive products: discussion paper based on the experience using a general Dutch prescription guideline. Prosthet Orthot Int 2011; 35(3): 310-7.
29. Jorgensen U, Melchiorsen H, Gottlieb AG, Hallas V, Nielsen CV. Using the International Classification of Functioning, Disability and Health (ICF) to describe the functioning of traumatised refugees. Torture 2010; 20(2): 57-75.
30. Leonardi M, Raggi A, Antozzi C, Confalonieri P, Maggi L, Cornelio F, Mantegazza R. Identification of international classification of functioning, disability and health relevant categories to describe functioning and disability of patients with myasthenia gravis. Disabil Rehabil 2009; 31(24): 2041-6.
31. Leach E, Cornwell P, Fleming J, Haines T. Patient centered goal-setting in a subacute rehabilitation setting. Disabil Rehabil 2010; 32(2): 159-72.
32. Schneidert M, Hurst R, Miller J, Ustun B. The role of environment in the International Classification of Functioning, Disability and Health (ICF). Disabil Rehabil 2003; 25(11-12): 588-95.
33. Boivin A, Currie K, Fervers B, Gracia J, James M, Marshall C, Sakala C, Sanger S, Strid J, Thomas V, van der Weijden T, Grol R, Burgers J. Patient and public involvement in clinical guidelines: international experiences and future perspectives. Qual Saf Health Care 2010; 19(5): e22.
34. Towle A, Godolphin W. Framework for teaching and learning informed shared decision making. BMJ 1999; 319(7212): 766-71.
35. Weston WW. Informed and shared decision-making: the crux of patient-centered care. CMAJ 2001; 165(4): 438-9.
36. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002; 324(7351): 1417.
37. Fung CH, Hays RD. Prospects and challenges in using patient-reported outcomes in clinical practice. Qual Life Res 2008; 17(10): 1297-302.
38. Valderas JM, Alonso J. Patient reported outcome measures: a model-based classification system for research and clinical practice. Qual Life Res 2008; 17(9): 1125-35.
39. Guyatt GH, Ferrans CE, Halyard MY, Revicki DA, Symonds TL, Varricchio CG, Kotzeva A, Valderas JM, Alonso J. Exploration of the value of health-related quality-of-life information from clinical research and into clinical practice. Mayo Clin Proc 2007; 82(10): 1229-39.
40. Au HJ, Ringash J, Brundage M, Palmer M, Richardson H, Meyer RM. Added value of health-related quality of life measurement in cancer clinical trials: the experience of the NCIC CTG. Expert Rev Pharmacoecon Outcomes Res 2010; 10(2): 119-28.
41. McKenna SP. Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. BMC Med 2011; 14;9:86.
42. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, Revicki DA, Symonds T, Parada A, Alonso J. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008; 17(2): 179-93.
43. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? Soc Sci Med 2005; 60(4): 833-43.
44. Wagner EH. Meeting the needs of chronically ill people. BMJ 2001; 323(7319): 945-6.
45. Bauer MS. A review of quantitative studies of adherence to mental health clinical practice guidelines. Harv Rev Psychiatry 2002; 10(3): 138-53.
46. Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. Br J Gen Pract 2007; 57(545): 971-8.
47. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, Whitty P, Eccles MP, Matowe L, Shirran L, Wensing M, Dijkstra R, Donaldson C. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6): iii-iv, 1-72.

48. Stewart RE, Vroegop S, Kamps GB, van der Werf GT, Meyboom-de Jong B. Factors influencing adherence to guidelines in general practice. Int J Technol Assess Health Care 2003; 19(3): 546-54.
49. Bensing J. Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine. Patient Educ Couns 2000; 39(1): 17-25.
50. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet 2005; 365(9453): 82-93.
51. Cranney M, Warren E, Barton S, Gardner K, Walley T. Why do GPs not implement evidence-based guidelines? A descriptive study. Fam Pract 2001; 18(4): 359-63.
52. Lugtenberg M, Burgers JS, Clancy C, Westert GP, Schneider EC. Current guidelines have limited applicability to patients with comorbid conditions: a systematic analysis of evidence-based guidelines. PloS one 2011; 6(10): e25987.
53. Cerniauskaite M, Quintas R, Boldt C, Raggi A, Cieza A, Bickenbach JE, Leonardi M. Systematic literature review on ICF from 2001 to 2009: its use, implementation and operationalisation. Disabil Rehabil 2011; 33(4): 281-309.
54. Lohr KN, Zebrack BJ. Using patient-reported outcomes in clinical practice: challenges and opportunities. Qual Life Res 2009; 18(1): 99-107.
55. Hahn EA, Cella D, Chassany O, Fairclough DL, Wong GY, Hays RD. Precision of health-related quality-of-life data compared with other clinical measures. Mayo Clin Proc 2007; 82(10): 1244-54.
56. Bausewein C, Simon ST, Benalia H, Downing J, Mwangi-Powell FN, Daveson BA, Harding R, Higginson IJ. Implementing patient reported outcome measures (PROMs) in palliative care--users' cry for help. Health Qual Life Outcomes 2011; 927.
57. Cano SJ, Hobart JC. The problem with health measurement. Patient Prefer Adherence 2011; 5279-90.
58. Brundage M, Bass B, Jolie R, Foley K. A knowledge translation challenge: clinical use of quality of life data from cancer clinical trials. Qual Life Res 2011; 20(7): 979-85.
59. Legare F, Ratte S, Stacey D, Kryworuchko J, Gravel K, Graham ID, Turcotte S. Interventions for improving the adoption of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2010(5): CD006732.
60. Gravel K, Legare F, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: a systematic review of health professionals' perceptions. Implement Sci 2006; 116.
61. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, Fervers B, Graham ID, Hanna SE, Makarski J. Development of the AGREE II, part 1: performance, usefulness and areas for improvement. CMAJ 2010; 182(10): 1045-52.
62. Collaboration A. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality and Safety in Health Care 2003; 12(1): 18-23.
63. Qaseem A, Forland F, Macbeth F, Ollenschlager G, Phillips S, van der Wees P. Guidelines International Network: Toward International Standards for Clinical Practice Guidelines. Ann Intern Med 2012; 156(7): 525-31.
64. Tong A, Lopez-Vargas P, Howell M, Phoon R, Johnson D, Campbell D, Walker RG, Craig JC. Consumer involvement in topic and outcome selection in the development of clinical practice guidelines. Health Expect 2011; 15(4): 410-23.
65. Uphoff EP, Wennekes L, Punt CJ, Grol RP, Wollersheim HC, Hermens RP, Ottevanger PB. Development of generic quality indicators for patient-centered cancer care by using a RAND modified Delphi method. Cancer Nurs 2012; 35(1): 29-37.
66. Yen L, Gillespie J, Rn YH, Kljakovic M, Anne Brien J, Jan S, Lehnbom E, Pearce-Brown C, Usherwood T. Health professionals, patients and chronic illness policy: a qualitative study. Health Expect 2011; 14(1): 10-20.
67. Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, Llewellyn-Thomas H, Lyddiatt A, Legare F, Thomson R. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011(10): CD001431.
68. van der Weijden T, Boivin A, Burgers J, Schunemann HJ, Elwyn G. Clinical practice guidelines and patient decision aids. An inevitable relationship. J Clin Epidemiol 2012; 65(6): 584-89.
69. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 1999; 318(7182): 527-30.

## Appendix 1 Members of t G-I-N Allied Health Steering Group

S. Bazin (SB), Chartered Society of Physiotherapy, London, UK. D. van Benthem (DvB), Dutch Association of Integrated Cancer Centres, Utrecht, The Netherlands. D. Dreesens (DD), Dutch Councis for Quality of Care, The Hague, The Netherlands. S. A. van Dulmen (SD), Radboud University Nijmegen Medical Centre, Scientific Institute for Quality of Healthcare, Nijmegen, The Netherlands. J. Gordon (JG), Royal College of Nursing, London, UK. S. Lukersmith (SL), Lukersmith & Assoc., Sydney, NSW, Australia. J. Muxlow (JM), School of Nursing, Dalhousie University Halifax, Nova Scotia, Canada. E. Santa Mina (ESM), Ryerson University Toronto, Toronto, ON, Canada. P. J. van der Wees (PW), Radboud University Nijmegen Medical Centre, Scientific Institute for Quality of Healthcare, Nijmegen, The Netherlands.

# 6

# Patient reported outcome measures (PROMs) for goal setting and outcome measurement in primary care physiotherapy: an explorative field study

**In press as**

van Dulmen SA, van der Wees PJ, Staal JB, Braspenning JCC, Nijhuis-van der Sanden MWG. Patient reported outcome measures (PROMs) for goalsetting and outcome measurement in primary care physiotherapy, an explorative field study. Physiotherapy (2016). http://dx.doi.org/10.1016/j.physio.2016.01.001

## Abstract

**Objectives** Routine use of patient reported outcome measures (PROMs) may provide an effective way of monitoring patient valued outcomes. In this study we explored (1) the current use of PROMs; (2) to what extent the goals correspond with the selected PROMs; (3) the health outcomes based on PROMs.
**Design** Observational clinical cohort study.
**Setting** Dutch primary care physiotherapy practices ($n$=43).
**Participants** Patients ($n$=299) with neck pain or low back pain.
**Main outcome measures** The number of PROMs used per patient were calculated. The International Classification of Functioning, Disability and Health was used to map the patients' goals and the percentages of PROMS selected that match the domains of the goals were calculated. Health outcomes were assessed using two approaches for estimating the minimal clinically important difference (MCID).
**Results** Repeated measurements with the Visual Analogue Scale, the Patient Specific Complaints questionnaire, the Quebec Back Pain Disability Scale, or the Neck Disability Index were completed by more than 60% of the patients. The PROMs used matched in 46% of the cases with goals for pain improvement, and in 43% with goals set at activity/participation level. The mean differences between baseline and follow up scores for all PROMs were statistically significant. Improvements of patients based on MCID varied from 57% to 90%.
**Conclusions** PROMs were used in the majority of the patients, showed improved health outcomes and fitted moderately with goals. The results of this study can be used for future research assessing the routine use of outcome measurements with PROMs.

## Introduction

The routine use of patient reported outcome measures (PROMs) is of increasing interest, because PROMs may provide an effective way of monitoring patient valued outcomes. The use of PROMs might facilitate goal setting with the patient in selecting health outcomes of the highest priority, and in determining what interventions are most likely to achieve those outcomes.[1,2] PROMs data provide an added value to other measurements, such as joint mobility, as PROMs capture information regarding aspects of health problems that are important to patients. Physiotherapist use goal setting for action planning, to set shared goals in consultation with the patient, and to examine to what extend the goals are achieved. PROMs may help the physiotherapist to identify goals, because patients are actively involved identifying their individual problems and goals which might increase their participation and satisfaction.[2-4] To interpret if goals have been achieved, it is important to compare PROMs scores before and after the intervention. Baseline, follow up scores, and mean change scores on the PROMs, or proportion of patients achieving a significant improvement based on the minimal clinically important difference (MCID) are frequently used metrics.[5] The MCID is defined as the minimal change that is important to the patient, and is considered an important parameter to enable a proper interpretation of change scores.[6]

Outcome measures at the patient level can be aggregated into population measures for feedback on the management of groups of patients. In quality systems this information is often presented using quality indicators for reflection to a certain benchmark. Recently, it has been advocated that PROMs should be integrated with care delivery as a useful quality measurement.[2,4,7] Therefore, we started to study the actual use of PROMs in daily physiotherapy practice.

Physiotherapists in the Netherlands are expected to be familiar with the use of PROMS as in Dutch physiotherapy guidelines many PROMs are recommended. However, little is known about the actual use of PROMs, the relation between goals and PROMs used and what the outcomes on the PROMs are. Therefore, our study aimed to improve insight into how PROMs are used in current clinical practice. The research questions were: 1) to what extent do physiotherapists apply PROMs in patients with neck and back pain; 2) to what extent do the domains of goals correspond with the PROMs selected; and 3) what are the actual health outcomes indicated by the PROMs when comparing baseline and follow-up measures.

# Methods

## Setting and participants

Data was collected from February till November 2010. Recruitment of physiotherapist working in primary care practices took place at an open meeting during the annual congress of physiotherapy in 2009. Upon enrolment, physiotherapists received an instruction guide and additional information about the study purposes. Physiotherapists ($n$=60) of 43 practices decided to participate in the study. Each participating practice formally provided written consent to participate and all patients were informed about the study and gave permission to use the data anonymously.

## Measures

The study included patients aged 18 years or older, who completed their intervention episode. These patients had been referred by a general practitioner or a medical specialist due to neck or low back problems, or used self-referral to consult a physiotherapist. In case of self-referral, patients were only included if they were diagnosed with neck pain or low back pain by the physiotherapists. All information regarding the treatment episode of each patient was recorded in an electronic health record (EHR). Patients were provided with a login code to complete pre- and post-treatment scores of PROMs online which were directly visible in the patient file. The results of the PROMs were visible in all phases of the diagnostics and treatment and facilitated thereby the clinical reasoning process. The physiotherapists identified goals based on the history taking (including the scores on the PROMs) and diagnosis and these goals were recorded on a pre-structured form based on shortlist of categories of the International Classification of Functioning, Disability and Health (ICF). The ICF provides a framework for understanding relationships between health components of functioning and disability, and contains five domains: a) body functions and structures; b) activities; c) participation; 4) environmental factors; and 5) personal factors.[3,8] In the EHR-system a shortlist of ICF categories was provided relevant for neck and low back problems to enable a more practical application of the ICF (see table 1). The ICF items included in the EHR were based on existing ICF core sets for low back pain[9], and modified by including items specific for neck pain and combining items for practical application.

PROMs selected were the Visual Analogue Scale (VAS) for pain measures; and for limitations in activities and participations the Patient Specific Complaints (PSC),[10] the Quebec Back Pain Disability Scale (QBPDS),[11] and the Neck Disability Index (NDI) were used.[12] For interpretation of change scores on the PROMs, cut off points for MCID were identified. The VAS scale measures intensity of pain on a continuous scale ranging from 0 mm (no pain) to 100 mm (worst pain imaginable) and the MCID is 15 mm.[13] The PSC is used for reporting problems with functions or activities.

Patients are asked to identify up to three important activities that they are having difficulty with or are unable to perform on a predefined list. Subsequently, they are asked to rate each specified activity on a scale with a range from 0 (able to perform activity without problems) to 100 mm (unable to perform activity). In this study the score for each participant's first activity was used. Based on an evaluation of the responsiveness of the PSC, cut off points for minimal improvement on the PSC have been estimated to range from 18 to 24 mm.[10] In present study we used a cut-off point for the MCID of 20 mm, based on research of Oostendorp et al.[14] The QBPDS is a 20-item questionnaire concerning activities of daily living in relation to low back pain. Responses on each item range from 0 (no disability) to 5 (total disability) with a total score ranging from 0 to 100[11] and the MCID is 20 points.[13] The NDI is a 10-item questionnaire to measure functional status of patients with neck pain.

For each item, answering option range from 0 (no disability) to 5 (total disability) resulting in a total range from 0-50 and the MCID is 7 points.[15]

## Statistical analyses

Descriptive statistics were used for calculating the number of episodes, patient characteristics, the use and the scores on the PROMs measurements. Baseline measures on the VAS and PSC, characteristics of the patients with neck pain and low back pain, and patients with a baseline and follow up measurement vs. without baseline and follow up measurement were compared using chi-squared statistics and unpaired t tests. Descriptive statistics and frequency tables were used to represent the extent of the match between the domain of the patients goal, and the domain captured by the PROM. If pain was selected as a goal, we calculated the frequency that the VAS was used and for goals on activities/participation level, we calculated frequencies of the PSC, NDI or QBPDS.

Because repeated measurements are clustered within a patient, the analyses were based on a linear mixed effect model with a random intercept and all other variables fixed. In order to control for patient characteristics[16-18], we included gender, age, chronic complaints (>12 weeks), and comorbidities (relevant secondary pathology, previous diseases, or medical treatments) as fixed effects.

There are several approaches in estimating MCID values due to different patient populations and different definitions of improvement on the anchor.[19] We used two methods suggested for calculating the proportion of patients who reached the threshold for a MCID. First, we assigned a dichotomous score for a clinically important improvement per PROM, based on an absolute MCID cut off point. For each instrument we determined a score 1=clinically important improvement, or 0=no clinically important improvement. Second we calculated a dichotomous score per PROM based on 30% improvement from baseline.[13]

**Table 1** Goals according to ICF domains, classified by the physiotherapist using the EHR framework ($n$=299 patients, missing $n$=2)

| **% of patients in which items of body functions and structure were selected** | | **% of patients in which items of activities were selected** | | **% of patients in which items of participation were selected** | |
|---|---|---|---|---|---|
| Respiratory / cardio-vascular system | 4.7 | Communication | 0 | Domestic life | 46 |
| Movement | 68 | Changing body position | 62 | Education / work / employment | 36 |
| Joints and bones | 67 | Maintaining body position | 9 | Community/ social/ civic life | 30 |
| Sleep/Global mental | 10 | Walking and moving | 27 | Interpersonal interactions | 3 |
| Neurological | 2 | Fine hand use | 18 | Other | 1 |
| Pain and sensory | 71 | carrying/moving/ handling objects | 40 | | |
| Muscle | 60 | Self-care | 5 | | |
| Other | 2 | Moving around using transportation | 9 | | |
| | | Other | 7 | | |
| **Total body functions and structures (% of patients)** | 846 (100) | **Total activities (% of patients)** | 551 (85) | **Total participation (% of patients)** | 344 (75) |
| Mean | 3 | % patients activities not selected | 15 | % patients participation not selected | 25 |
| Median | 3 | Mean | 1.9 | Mean | 1.2 |
| Range | 0-7 | Median | 2 | Median | 1 |
| | | Range | 0-6 | Range | 0-4 |

| | Environmental factors | | | |
|---|---|---|---|---|
| | **% patients in which items of external factors were selected** | | **% patients in which items of personal factors were selected** | |
| | Natural environment | 2 | Pain perception / pain behaviour | 13 |
| | Social support | 2 | Coping | 5 |
| | Working environment (technology /services) | 16 | Own thoughts about the health problem | 14 |
| | Other | 16 | Lifestyle | 30 |
| | | | Social economic situation | 1 |
| | | | Stress | 14 |
| | | | Expectations of recovery | 10 |
| | | | Other | 1 |
| | | | | |
| | **Total external factors (% of patients)** | 71 (28) | **Total personal factors (% of patients)** | 199 (44) |
| | % patients external factors not selected | 72 | % patients personal factors not selected | 56 |
| | Mean | 0.2 | Mean | 0.7 |
| | Median | 0 | Median | 0 |
| | Range | 0-3 | Range | 0-5 |

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 20.0 (IBM Corp, Armonk, New York).

# Results

## Responses and characteristics of participants

A total number of 60 physiotherapists (19% female) participated; 23% of the physiotherapists were aged <39 years, 27% between 40 and 49 years and 52% of the physiotherapist were 50 years or older. The practices were all private practice clinics, 31% of the practices were solo practices and on average 3.3 therapists worked per practice. The participating physiotherapists comprised a representative sample when compared with national reference data.[20] Only the percentage of women differed: 19% women in our study compared to 56% women in reference data.

A total of 561 patients with neck pain or low back pain were entered in the EHRs, of which 299 met the inclusion criteria. The excluded patients had uncompleted treatment episodes or were aged under 18 years. The patient characteristics gender, age, and the ratio of patients with neck and low back pain comprised a representative sample when compared to national reference data[21], see table 2. In 7% of the patients with neck pain, the physiotherapists also indicated a problem with the lower back. In 5% of the patients with low back pain a problem with the neck as a relevant secondary pathology was indicated. Characteristics of patients with neck pain and low back pain did not differ significantly, except for gender; more female patients with neck pain were treated by the physiotherapist than female patients with low back pain. Patients without a PROM measurement were significantly younger than patients with a PROM measurement (see table 2).

## PROM measurement

One baseline measurement was completed by 89 patients with neck pain (89/128, 70%), and a repeated measurement was completed by 78 patients (78/128, 61%). One baseline measurement was completed by 116 patients with low back pain (116/171, 68%), and a repeated measurement was completed by 108 patients (108/171, 63%). In 76 patients (76/299, 25%) PROMs were not used. The proportion of the different PROMs used is described in table 3.

**Table 2** Patient characteristics

| | low back pain (n=171) | Neck pain (n=128) | P-value | Patients with baseline or follow up measurement (n=223) | Patients without measurement (n=76) | P-value |
|---|---|---|---|---|---|---|
| Gender Female% | 54 | 67 | .02* | 61 | 54 | .52 |
| Mean age (SD) % | 50 (16) | 49 (15) | .51 | 51 (15) | 46 (16) | .00* |
| Chronic complaints (>3 months) % | 52 | 58 | .21 | 55 | 51 | .51 |
| Comorbidities yes % | 67 | 69 | .31 | 65 | 62 | .62 |

* Statistically significant ($P < 0.05$)

**Table 3** he proportion of measurements used

| | Patients with neck pain (n=128) | | | | | Patients with low back pain (n=171) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | NDI (%) | PSC (%) | VAS (%) | Total measures | N Patients (%) | QBPDS (%) | PSC (%) | VAS (%) | Total measures | N Patients (%) |
| **Baseline measurement** | 47 (37) | 19 (15) | 57 (45) | 123 | 89 (70) | 55 (32) | 37 (22) | 78 (46) | 170 | 116 (68) |
| **Follow up measurement** | 43 (34) | 18 (14) | 50 (39) | 111 | 78 (61) | 48 (28) | 34 (20) | 74 (43) | 156 | 108 (63) |

QBPDS: Quebec back pain disability scale, NDI: Neck disability index, PSC: Patient specific complaints, VAS: Visual analogue scale

## Goals

Physiotherapists recorded a mean of 7 goals (range 0-19), see table 1. Figure 1 shows to what extend the selected domains of the goals correspond with the PROMs. In 211 patients pain was indicated as goal, for which in 96 patients a VAS was used (96/211, 45%). For 238 patients goals were selected at the activity/participation level, and 102 patients completed the NDI, QBPDS or PSC (102/211, 43%) see figure 1. The goals of 255 patients included both pain and activity/participation domains of which 93 patients (93/255, 36%) completed a VAS and a NDI, QBPDS or PSC. The open text fields of the category 'other' is described in Appendix 1.

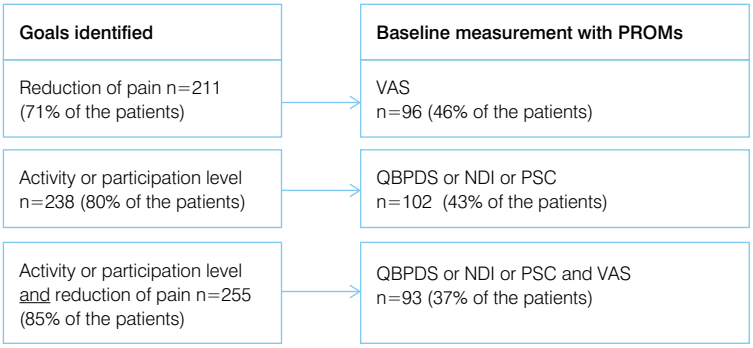


**Figure 1** Domains of goals identified and PROMs selected

## Outcomes

Patient characteristics did not differ significantly between the patients with a baseline and follow up measurement and patients without a baseline and follow up measurement. Table 4 describes the uncorrected scores of the PROMS and the percentage of patients who improved based on the MCID. The differences between baseline and follow up scores for all PROMs were statistically significant for both patient groups. The mixed model analysis showed a statistically significant effect of change score on the outcomes of all PROMs (supplementary table 5).

**Table 4** Baseline, follow-up and change scores and percentage of patients improved based on MCID

| | Mean baseline score | SD | N | Mean follow up score | SD | N | Mean change score (uncorrected) | 95% CI | *MCID* absolute improvement points | *MCID* 30% improvement |
|---|---|---|---|---|---|---|---|---|---|---|
| **QBPDS** | 38.7 | 20.1 | 55 | 15.3 | 15.0 | 48 | -24.4* | -18.6, -30.1 | 67 | 83 |
| **NDI** | 15.6 | 6.5 | 49 | 6.3 | 6.5 | 45 | -9.2* | -7.0, -11.4 | 57 | 82 |
| **PSC** | 54.7 | 19.6 | 56 | 18.1 | 21.4 | 52 | -36.4* | -30.4, -42.5 | 81 | 85 |
| **VAS** | 44.7 | 18.2 | 135 | 12.0 | 18.3 | 124 | -32.7* | -28.5, -36.7 | 80 | 90 |

QBPDS: Quebec back pain disability scale (MCID ≥ 20 points); NDI: Neck disability index (MCID ≥ 7 points)
PSC: Patient specific complaints. In this study only the first activity was used.e (MCID ≥ 20 mm) first
VAS: Visual analogue scale (MCID ≥ 15 mm); 95CI: 95% confidence interval; * Statistically significant ($P < 0.05$)

## Discussion

Our study showed in general that PROMs were used for baseline and follow up measurements in the majority of patients. The selected domains of the goals matched the domain of the PROMs moderately. In 46% of the patients with a goal on pain reduction a VAS was used, whereas in 43% of the patients with a goal on activity or participation level a QBPDS, NDI or PSC was used.

The PROMs were used for more than 60% of the patients, which is high compared to other studies[22]. This might be explained by the fact that our study was conducted with a small self-selected sample of physiotherapists who might be more motivated to use PROMs, and the fact that patients could complete the PROMs questionnaires online.

The high number of goals is comparable with other studies.[23,24] This may be explained by the overlap between the different items, which makes it difficult to value the most important goals and to assess which PROM is most appropriate. The ICF list we used in our study did not rank the selected items and did not provide an explicit short term or long term goal. Further specification of the most important goals is necessary to indicate which PROM or other outcome measure should be used. Another reason for the discrepancy between the selected goals and PROMs may be due to the fact that the ICF domains represent more categories than is covered by the selected PROMs. The ICF model includes environmental and personal factors as important modifying variables in the development of patients' goals and interventions.[25] For example, specific educational interventions in case of avoidance or pain coping may be relevant, such as explanation and instructions focused on behavioural change. Of further consideration is the selection of the appropriate items used for goal setting. We did not perform a systematic approach for linking all items of the PROMs to ICF categories, nor did we examine in what way goals are selected. Goals may also be selected based on interaction with the patients without using PROMs.

The selected PROMs showed overall improved health outcomes that were clinically significant, with levels of improvement that were comparable with other studies[26,27].

Our study demonstrated differences in the percentage of patients that improved based on different MCID cut off values. All PROMs showed floor effects, indicating that baseline scores of some patients were lower than the absolute MCID cut off points. The empirical evidence on optimal methods for estimating the MCID is heterogeneous and the cut off points are varying.[28] Although the MCID reflects a change on group level, this does not necessarily match the proposed change score at individual level, as it is known that PROMs may be less reliable at the patient level.[29] Future research should focus on how the scores of the PROMs are embedded in the goal setting process and in the conversation with the patient. More research needs

to be undertaken to identify clinically important improvements for the diverse patient groups.

Physical therapists should be encouraged to collect routine PROMs data. In Dutch physiotherapy guidelines PROMs are recommended although their implementation is limited. Strategies that focus on implementation of guidelines, and how PROMs can assist physiotherapists in their clinical reasoning process could improve the use of PROMs.

The strength of this study is that the results are based on observational cohort data with all types of neck pain and low back pain. Patients comprised a representative sample of the population visiting a Dutch physiotherapy practice. This indicates that PROMs are potentially useful for patients with neck pain or low back pain for outcome comparison at group level. A few limitations need to be addressed. This was an explorative field study with a small sample. We have included all type of patients with low back and neck pain and we accounted for some confounding variables only, and we might account for other patient characteristics in future studies. Moreover we did not examine how PROMs are used in the goal setting process and how outcomes are related to the intervention.

The results of this study can be used for future research assessing the routine use of outcome measurements with PROMs in identifying clinical relevant improvement of PROMs based on observational cohort data. Additional research may provide insight in how PROMs are used in the goal setting proces. The results of this study can also be used for identifying an ICF core set which match the PROMs used and which are feasible for use in clinical practice.

# References


1. Reuben DB, Tinetti ME. Goal-oriented patient care - an alternative health outcomes paradigm. N Engl J Med 2012; 366(9): 777-9.
2. Van der Wees P, Nijhuis-van der Sanden M, Ayanian JZ, Black N et al. Integrating the use of patient-reported outcomes for both clinical practice and for performance measurement: experts' views from three countries. Milbank Q 2014; 92(4): 754-75.
3. van Dulmen SA, Lukersmith S, Muxlow J, Santa Mina E et al. Supporting a person-centred approach in clinical guidelines. A position paper of the Allied Health Community - Guidelines International Network (G-I-N). Health Expect 2015; 18(5): 154-58.
4. Kyte DG, Calvert M, van der Wees PJ, Ten Hove R et al. An introduction to patient-reported outcome measures (PROMs) in physiotherapy. Physiotherapy 2014; 101(2): 119-25.
5. Hildon Z, Neuburger J, Allwood D, van der Meulen J et al. Clinicians' and patients' views of metrics of change derived from patient reported outcome measures (PROMs) for comparing providers' performance of surgery. BMC Health Serv Res 2012; 12171.
6. Copay AG, Subach BR, Glassman SD, Polly DW et al. Understanding the minimum clinically important difference: a review of concepts and methods. Spine Journal 2007; 7(5): 541-6.
7. McGlynn EA, Schneider EC, Kerr EA. Reimagining quality measurement. N Engl J Med 2014; 371(23): 2150-3.
8. Escorpizo R, Bemis-Dougherty A. Introduction to Special Issue: A Review of the International Classification of Functioning, Disability and Health and Physical Therapy over the Years. Physiother Res Int 2015; 20: 200–9.
9. Cieza A, Stucki G, Weigl M, Disler P et al. ICF Core Sets for low back pain. Journal of Rehabilitation Medicine 2004; (44 Suppl): 69-74.
10. Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996; 65(1): 71-6.
11. Schoppink LE, van Tulder MW, Koes BW, Beurskens SA, de Bie RA. Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. Phys Ther 1996; 76(3): 268-75.
12. Jorritsma W, de Vries GE, Dijkstra PU, Geertzen JH, Reneman MF. Neck Pain and Disability Scale and Neck Disability Index: validity of Dutch language versions. European Spine Journal 2012; 21(1): 93-100.
13. Ostelo RWJG, Deyo RA, Stratford P, Waddell G et al. Interpreting change scores for pain and functional status in low back pain - Towards international consensus regarding minimal important change. Spine 2008; 33(1): 90-94.
14. Oostendorp RA, Rutten GM, Dommerholt J, Nijhuis-van der Sanden MW, et al. Guideline-based development and practice test of quality indicators for physiotherapy care in patients with neck pain. J Eval Clin Pract 2013; 19(6): 1044-53.
15. Young IA, Cleland JA, Michener LA, Brown C. Reliability, construct validity, and responsiveness of the neck disability index, patient-specific functional scale, and numeric pain rating scale in patients with cervical radiculopathy. American Journal of Physical Medicine & Rehabilitation 2010; 89(10): 831-9.
16. Walton DM, Carroll LJ, Kasch H, Sterling M et al. An Overview of Systematic Reviews on Prognostic Factors in Neck Pain: Results from the International Collaboration on Neck Pain (ICON) Project. The open orthopaedics journal 2013; 7494-505.
17. Pincus T, Santos R, Breen A, Burton AK, Underwood M, Multinational Musculoskeletal Inception Cohort Study C. A review and proposal for a core set of factors for prospective cohorts in low back pain: a consensus statement. Arthritis and Rheumatism 2008; 59(1): 14-24.
18. Bekkering GE, Hendriks HJ, van Tulder MW, Knol DL et al. Prognostic factors for low back pain in patients referred for physiotherapy: comparing outcomes and varying modeling techniques. Spine 2005; 30(16): 1881-6.
19. Dworkin RH, Turk DC, Wyrwich KW, Beaton D, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008; 9(2): 105-21.
20. Nivel. Beroepen in de gezondheidszorg [Professions in healthcare]. Utrecht, the Netherlands: Nivel.

21. Kooijman MK, Verberne LDM, Barten JA, Leemrijse CJ et al. NIVEL Zorgregistraties eerste lijn [NIVEL health care registrations primary care]. 2013. http://www.nivel.nl/NZR/fysiotherapie [retrieved 11.12.14].
22. Swinkels RA, van Peppen RP, Wittink H, Custers JW, Beurskens AJ. Current use and barriers and facilitators for implementation of standardised measures in physical therapy in the Netherlands. BMC Musculoskeletal Disorders 2011; 12106.
23. Huber EO, Tobler A, Gloor-Juzi T, Grill E, Gubler-Gut B. The ICF as a way to specify goals and to assess the outcome of physiotherapeutic interventions in the acute hospital. Journal of Rehabilitation Medicine 2011; 43(2): 174-7.
24. Finger ME, Selb M, De Bie R, Escorpizo R. Using the International Classification of Functioning, Disability and Health in Physiotherapy in Multidisciplinary Vocational Rehabilitation: A Case Study of Low Back Pain. Physiother Res Int 2015;20:231–41.
25. Helgeson K, Smith AR, Jr. Process for applying the international classification of functioning, disability and health model to a patient with patellar dislocation. Phys Ther 2008; 88(8): 956-64.
26. Pool JJ, Ostelo RW, Knol DL, Vlaeyen JW et al. Is a behavioral graded activity program more effective than manual therapy in patients with subacute neck pain? Results of a randomized clinical trial. Spine 2010; 35(10): 1017-24.
27. Artus M, van der Windt D, Jordan KP, Croft PR. The clinical course of low back pain: a meta-analysis comparing outcomes in randomised clinical trials (RCTs) and observational studies. BMC Musculoskeletal Disorders 2014; 15:68.
28. Schuller W, Ostelo RW, Janssen R, de Vet HC. The influence of study population and definition of improvement on the smallest detectable change and the minimal important change of the neck disability index. Health and Quality of Life Outcomes 2014; 12:53.
29. de Vet HCW, Terwee CB, Mokkink LB, Knol D, L. Measurement in medicine: A practical guide. Cambridge: Cambridge University Press, 2011.

**Table 5** Results of the linear mixed model analysis

| | **QBPDS** (0 -100) | | | **NDI** (0-50) | | | **PSC** (0-100) | | | **VAS** (0-100) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ***β*** | ***P*** | **95%CI** | ***β*** | ***P*** | **95%CI** | ***β*** | ***P*** | **95%CI** | ***β*** | ***P*** | **95%CI** |
| **Intercept** | 52.9 | 0.00* | 36.7, 69.2 | 17.0 | 0.00* | 10.1, 23.8 | 68.2 | 0.00* | 47.2, 89.2 | 51.0 | 0.00* | 40.0, 61.9 |
| **Change score in time (follow up – baseline)** | -24.9 | 0.00* | -30.6, -19.3 | -9.3 | 0.00* | -11.6 -7.1 | -36.5 | 0.00* | -42.6, -30.4 | -32.8 | 0.00* | -36.9, -28.7 |
| **Gender Female** | 5.9 | 0.13 | -1.8, 13.6 | 1.8 | 0.30 | -1.7, 5.2 | -1.2 | 0.79 | -10.8, 8.3 | 0.3 | 0.91 | -5.3, 6.0 |
| **Age** | -0.1 | 0.42 | -0.4, 0.2 | -0.0 | 0.98 | -0.1, 0.1 | -0.14 | 0.40 | -0.5, 0.2 | -0.1 | 0.44 | -0.3, 0,1 |
| **Chronic complaints (>3 months)** | 4.0 | 0.32 | -3.9, 11.9 | -1.5 | 0.36 | -4.8, 1.8 | 2.7 | 0.59 | -7.0, 12.3 | -2.1 | 0.43 | -7.5, 3.1 |
| **Comorbidities (yes)** | 14.9 | 0.02* | -5.8, 24.1 | 3.5 | 0.04* | -0.2, 6.7 | 10.6 | 0.04* | -0.44, 20.8 | 9.2 | 0.02* | -3.5, 14.9 |

QBPDS: Quebec Back Pain Disability Scale
NDI: Neck Disability Index
PSC: Patient Specific Complaints
VAS: Visual Analogue Scale
95CI: 95% confidence interval
*$p<0.05$

## Appendix 1 Category 'other' of goals according to ICF domains

**1. Body functions and structure (n=5)**

- Recovery of extension and lateral flexion of the lumbar spine
- Recovery of mobility of the lumbar spine and strengthening of the musculature
- Improvement of mobility and capacity
- Relaxation
- Segmental stability

**2. Activities (n=21) (n=1 not specified)**

- Reducing kyphotic postures
- Improving the mobility of the neck when working above the head
- Bending over (n=2)
- Cervical rotation
- reducing sleeping hands
- Relaxation and sleep
- Transfer to bed and kneeling
- Playing sports
- Playing basketball
- Standing (n=2)
- Typing
- Sitting, standing, and bending over
- Sitting, standing, strolling
- Sitting posture (n=3)
- Working activities
- Specific neck positions

**3. Participation (n=4) (n=1 not specified)**

- Working in the garden
- Volunteering
- General more physically active

**4. Personal factors (n=9)**

- Advise on working position/activities
- Advise on working position and driving a car
- Swimming advice
- Assisting gym
- Advise to an optician
- Volunteering

- Stress at home
- Limited skills in Dutch language
- General more physically active

**5. External factors (n=3)**

- Assertiveness in relation to the working environment
- Learn to improve muscle function in a responsible and controlled way
- Assist with change of profession

# 7

# General discussion

# Introduction

The studies in this thesis were focused on identifying determinants for improving the quality of physical therapy care. Several studies listed knowledge, clinical reasoning, and treatment strategies that are geared towards the patient's needs and preferences, as well as monitoring of changes in health status as important components of patient-centered quality care.[1-3] Behavioral routines on the part of physical therapists, such as reflection and feedback, are key elements in improving the quality of care.[3,4] In reflecting on the results of this thesis, we will describe these elements more extensively, using two cases from the patient safety study in chapter 2 as examples.

# Results

## Awareness: an important element in clinical reasoning

### *Clinical reasoning and patient safety*

Lack of knowledge and errors in clinical reasoning are associated with patient safety incidents (see chapter 2). Box 1 presents a case history that illustrates unawareness on the part of a healthcare provider concerning the provision of sub-optimal care. Analysis of this case reveals that the patient's inability to recover was due to a incomplete diagnosis as consequence of the healthcare provider's insufficient knowledge, the lack of monitoring and evaluation of the patient's health status, and/or the lack of adequate response from the physical therapist (and other healthcare providers). In this case, the therapist continued extensive treatment despite the patient's lack of progress.

It is not clear whether the physical therapist was unaware of the deviating course of symptoms, or whether he did not adequately respond to the symptoms. Baseline measurements on muscle strength, balance and knee stability, pain, and functional limitations were not registered, and changes in health outcomes were impossible to compare. The increase in symptoms was not shared with the orthopedist, and the early diagnosis was not re-evaluated, although the signs and symptoms did not reflect the expectations for a patient who had undergone a 6-week casting period and was receiving professional treatment for a patellar luxation. Objective measurements, systematic reports in the health record and patient-reported outcomes would have been helpful in identifying this abnormal course at an earlier stage and may have led to consideration of the possibility that the patient was suffering from more complications than just a patellar luxation.

Our safety study in chapter 2 showed that causes of incidents were mainly related to knowledge-based behavior, communication with other healthcare providers, clinical

**Box 1** Mrs. M.

Woman, 40 years old

Diagnosis: patellar luxation

History and diagnostics: Mrs. M. was involved in a car accident, resulting in a patellar luxation. After treatment in the hospital, she was put in a cast for six weeks. Once the cast was removed, her physical therapist's intervention focused on improving her joint mobility and stability.

Interventions (80 treatments) consisted of exercise therapy aimed at improving stability and mobility of the knee, and improving muscle strength. The patient was treated daily for more than two months while symptoms increased: instability, inflammation, problems with the knee "giving way", and two incidents of falling. The patient's health record contains very little information on the interventions, or any changes in health status.

Four months later, a MRI showed a rupture of the anterior and posterior cruciate ligaments and the medial collateral ligaments.

reasoning, and inadequate monitoring. Allied healthcare professionals were unaware of their unsafe behavior, the harmful results for patients, and their own level of competence. This is due to the fact that they attributed the risk of unsafe care to patient-related or environment-related characteristics, rather than reflecting on their own professional behavior. This unawareness was likely the reason for the absence of reported incidents in our safety study.[5] Research has shown that healthcare professionals have a limited ability to accurately assess their own level of competence and lack (risk) awareness, as they do not recognize any safety problems in their current approach to work.[6-8] This is illustrated in case 1 (see box 1), where the physical therapist did not respond to an increase in symptoms and the abnormal course of symptoms. Diagnostic errors (and the failure to correct them) are related to problems in clinical reasoning skills, the physical therapists' insufficient knowledge about normal, expected recovery patterns, poor record keeping, organizational difficulties or ineffective communication between the healthcare provider and the patient. Diagnostic errors may also lead to inadequate interventions that yield insufficient results, or cause adverse events and malpractice complaints. This is illustrated in our patient safety study in physical therapy, but also applies to other healthcare providers.[9-11]

Greater awareness of errors in the reasoning process will help physical therapists to improve the quality of care and reduce adverse events.[12] On the organizational

level, regular discussions about medical incidents, in informal contexts as well as in structurally organized practice meetings, may increase risk awareness, openness and critical clinical reasoning. Physical therapists need to develop a clear picture of what constitutes a safety problem or incident, and of what to do to ensure safety. This process is one that requires integration of the logic of risk management and of clinical practice.[13] An open culture is needed to raise awareness. It is also a prerequisite for conducting activities that require openness and trust, such as reporting.[14]

### *Awareness of strengths and weaknesses*

Peer assessment leads to an increase in knowledge and guideline-consistent clinical reasoning, and is an effective method for enhancing awareness of strengths and weaknesses and reflective practice (see chapters 3 and 4).

The peer assessment groups showed a significantly greater awareness of their professional performance. The critical success factor for peer assessment can be assigned to the structured and performance-based design, which enhances clinical reflection and awareness of clinical performance. Reflection on decision-making is a powerful tool in developing clinical reasoning skills and is an important method of learning in practice.[12,15,16] The role of awareness was confirmed by the results of another implementation strategy for a Dutch guideline for physical therapist management of upper extremity conditions.[8] In this randomized controlled trial, which was conducted by Maas et al., a similar implementation strategy showed that peer assessment was significantly more effective in improving guideline adherence as compared to case-based discussions.

An important element of the peer assessment strategy is reasoning aloud, a process that makes physical therapists more aware of their strengths and weaknesses. Reasoning aloud also enabled peers to provide feedback regarding knowledge and perceptions. Moreover, this activity served repeatedly to challenge physical therapists to reproduce and apply newly acquired knowledge, thereby guiding self-directed change toward personal learning. The clinical reasoning process often takes place unconsciously, depending on the physical therapists' familiarity with the problem encountered. In the peer assessment groups, participants need to make the transfer from implicit reasoning to explicit reasoning, and from intentional behavior to observable behavior. This strategy makes physical therapists aware of their thinking process and of their biases, which in turn will help them to detect and correct their own potential reasoning flaws.

Errors in clinical reasoning are associated with hypothesis generation, clinical information collection, lack of knowledge, or interaction with the patient.[11] Errors are commonly associated with habits of thinking and practice (intuitive thinking), which are a potential risk in themselves to the strategy of pattern recognition.[15] Physical therapists in the peer assessment groups were repeatedly challenged to provide

insight into their own perceptions of the patients' needs. They were also encouraged to formulate their reasoning explicitly and reflect on it. In that process, they were guided by questions from colleagues, such as: why am I doing this? Or, why am I making this decision? This reflection may have strengthened awareness of deficiencies in knowledge of signs and symptoms, the course of development typically expected of various conditions and the relationship between diagnostic findings and therapeutic choices.

It is known that the clinical reasoning of experts in familiar situations frequently lacks an explicitly continuous process of hypothesis formulation and evaluation. It is rapid, automatic, and often, non-verbal. Easy cases are solved by pattern recognition and by direct association of the data with diagnostic classifications and interventions. Difficult cases require systematic hypothesis generation and testing. Whether a problem is easy or difficult will depend in part on the knowledge, experience, and communications skills of the therapist who is trying to solve it.[15] In optimal patient-centered care, physical therapists will reflect continually on their working hypotheses and the effects of their interventions to "validate" their clinical patterns and procedural knowledge.

Reflection on work routines is essential, as it serves to guide continued hypotheses generation and evaluation. It is through professional education and clinical experience, that physical therapists can identify the categories of information that they find particularly useful for problem identification and management decisions. Beyond these routines, specific inquiries and tests are tailored to each patient's unique presentation. This cognitive activity is called "hypothesis testing".[15] The resulting data are then interpreted and modified, after which new hypotheses are considered. Through a process of evaluating the patients' understanding of and feelings about their problems, as well as through explanation, reassurance and shared decision making, patients and their therapists develop an evolving understanding of the problem and its management.

## Supporting clinical reasoning

In clinical practice, physical therapists can use existing models for hypothesis testing and for taking account of patients' needs and expectations. The International Classification of Functioning (ICF) is a valuable tool for identifying limitations in activity and participation, and influencing personal or environmental factors as factors relating to the predisposition, development and maintenance of the problem. In the ICF, patients' health conditions can be seen both as aspects that influence and that are influenced by their body functions and structures, their capacity and performance of functional activities of life and their subsequent ability to participate in their family life, work and leisure. In addition, environmental or personal factors are perceived to have positive - or negative - impacts on patient health conditions. There

is a growing and encouraging amount of literature, where the ICF has been demonstrated to be useful in physical therapy.[17] Physical therapists can employ the ICF to evaluate their patients over time, using the ICF domains as a guide. To facilitate ICF use in practice, several practitioners have proposed conceptual models, such as ICF core sets, and case examples that utilize the ICF as a basis for decision-making.[18-21] An ICF core set is a selection of ICF categories from the full ICF classification that are considered most relevant for describing the functioning of a person with a specific health condition, or in a specific healthcare context.

Guidelines are valuable tools that can support physical therapists by providing information on the steps in the clinical reasoning process. Guidelines should be seen as an aid to physical therapists and patients weighing pros and cons in decision-making. However, when reflection and clinical judgment are lacking, one potential pitfall to using guidelines is that they could be employed as a protocol.

Clinical reasoning may also be enhanced by a tool that is helpful in setting priorities in physical therapy diagnoses and treatment. The Hypothesis Orientated Algorithm for Clinicians (HOAC) offers a conceptual, patient-centered framework for physical therapists with an algorithm for clinical reasoning that describes all of the steps for making appropriate decisions about the treatment of patients.[22] The HOAC focuses on actively involving patients in the decision-making process. Clinical reasoning, according to the HOAC algorithm, uses an approach in which hypotheses about the causes and consequences of the patient's problems are listed based on the anamnesis. These hypotheses are tested, using observations and measurement instruments that ensure that the treatment addresses the underlying causes of the most restrictive disabilities and limited areas of participation. The therapist then establishes testing criteria to evaluate the outcome of the intervention.

## Monitoring health outcomes

Outcome measures facilitate the dialogue between patients and physical therapists about what is important to them and about the relevance and effectiveness of the intervention (see chapter 5). PROMs are used for baseline and follow-up measurements in the majority of patients.

An important element of the clinical reasoning and goal setting process is that of determining outcomes in collaboration with the patient.[23] It is envisioned that PROMs can assist physical therapists in their clinical reasoning process for diagnostics and treatment, with a specific focus on the patient's perspective.[24] The use of PROMs in setting goals and evaluating the course and outcome of the treatment provides physical therapists with information relevant to their patients' progress. The interaction between the physical therapist and the patient in evaluating the outcomes of the PROMs may lead to alternative hypotheses, or to the consideration of other interventions if the expected results are not achieved. PROMs could, therefore,

improve patient outcomes by enabling physical therapists to detect and treat problems that may have been missed previously.[25]

**Box 2** Mrs. A.

Woman, age 37 years old

Diagnosis by general practitioner: Multiple sclerosis (> 5 years).

History and diagnostics: Mrs. A. suffered from pain in the lower back and neck, and a loss of muscle power resulting in an increasing inability to perform daily activities. Her neck mobility had decreased and there was hypertonia.

The treatment plan was aimed at reducing pain and improving spinal mobility. The baseline measurement was performed with the Patient Specific Complaints questionnaire (PSC).* According to the patient, the most important activities requiring improvement were: 1) walking; 2) rising from a chair and 3) lying in bed.

Interventions (20 treatments) consisted of passive techniques, such as massage and spinal mobilization. Mrs. A fell twice during the treatment period. The interventions failed to bring about any reduction in pain or improvements in Mrs. A's activities. Nonetheless, no actions were reported in her treatment log concerning any adjustments to the intervention strategy.

* The PSC is used for reporting problems with functions, or activities. Patients are asked to identify up to three important activities that they find difficult or are unable to perform on a predefined list. Subsequently, they are asked to rate each specified activity on a scale.

The case illustrated in box 2 is an example in which the intervention was not effective, not patient centered and not safe. The patient's request for help came from a desire to suffer less pain and improve in her ability to engage in activities. Better outcomes could probably have been achieved if the following hypotheses had been established in the diagnostic phase: 1) Multiple Sclerosis leads to a loss of muscle strength and coordination problems, and 2) the neck/low back pain are a result of overloading. It is unclear what the hypotheses were of the physical therapist in the diagnostic and treatment phases. The interventions only focused locally on body functions (mobility of the spine) and structures (muscles) of the neck and low back, while the request for help stemmed from the patient's limitations in activities and participation (as measured by the PSC questionnaire). The patient's care was unsafe and inadequate due to various factors, including: the exclusive focus on passive strategies; (which is ineffective as the sole treatment for patients with back pain):[26] the lack of attention for the patient's daily functioning or muscle core set: the failure to adapt the treatment

after two falling incidents; and the patient's deprivation of more appropriate care resulting from unnecessarily long intervention periods. Logically speaking, the treatment goals should also have addressed the symptoms related to Multiple Sclerosis, such as muscle strength, balance and physical activity, as exercise therapy is known to be effective in improving activities and participation, as well as the quality of life.[27] It is unclear how the physical therapist used the results of the PSC for diagnostics, in setting goals, and in interacting with the patient. Repeated measurements with the PSC questionnaire could have been helpful in assessing the effectiveness of current treatments. In a case where follow ups reveal that the outcomes on the PROMs have remained unchanged or have worsened, it is appropriate for the physical therapist to discuss the results with the patient and possibly to change the hypothesis or treatment plan, or even to refer him/her to a neurologist. The health domain(s) that are important for the patient selection, e.g. pain, activities, social functioning, should guide the selection of the PROMs. In this case, the PSC questionnaire was a suitable instrument. Additionally, an outcome measurement for pain, such as the Numeric Pain Rating Scale (NPRS), could be used to monitor the change in pain experience. Case 2 is illustrative of the study results outlined in chapter 6. In that study, we found that the selected domains of the goals matched the domain of the PROMs to a moderate extent. Although PROMs have the potential to facilitate shared identification of goals and priorities between health professional and patients, this is not yet common practice. Optimal use of collected data on the process and outcomes of care require solid technological support for collecting and sharing information between patients and physical therapists on the baseline health status, goals, course of further development, and outcomes of interventions.[28]

## Patient records

In several studies in this thesis, we rely on the reports in the patient records of the physical therapists. The study on patient safety (chapter 2) revealed that the poor quality of the patient records hindered the detection and assessment of incidents. It is hard to obtain information about all of the contributing factors in the care provided by reading patient records. This hinders the assessment of the clinical reasoning process and interaction with the patient. This was also difficult in our study on the use of PROMs for goal setting and outcome measurement (chapter 6). We could not examine how PROMs are used in the goal-setting process and how outcomes are related to the intervention.

Analysis of current care requires systematic recording of care. Very little effort is made to clarify implicit reasoning in the patient record, which makes the assessment of clinical reasoning sometimes difficult to interpret. On the other hand, record keeping may be a way to reflect explicitly on the physical therapist's thoughts. Process and outcome variables are needed to provide insight into the quality and

safety of care. Quality indicators can be used to assess the structure, process and outcome of actual care, in order to examine the gap between actual care and intended care. Future research is needed to identify core measures that best reflect the quality and outcomes of the care delivered, as such measures would be useful to different stakeholders, such as patients, healthcare providers, payers, and researchers.

## Implications

The task of improving patient-centered quality in healthcare requires action at the level of clinical guidelines, at the professional level and at the patient level.

### *1. Implications for guideline development and guideline implementation*

Clinical guidelines have the potential to stimulate patient-centered healthcare, although effective involvement of patients and inclusion of their preferences in guideline development varies. [29-32] In the Netherlands, several initiatives have been developed which will further stimulate patient-centered care in guideline development and implementation. A recent initiative is the development of a blueprint entitled "Patient Participation in Guideline Development," which is a tool that guideline developers can use to incorporate the patient's perspective effectively and efficiently in (evidence-based) guidelines.[33] In 2013, the HARING toolbox (Guidance for Guidelines in Dutch Healthcare) was published in order to support the development and implementation of guidelines. The toolbox includes an implementation checklist, which is intended as an aid for guideline groups in drawing up assessments of potential barriers to the dissemination and implementation of guidelines.[34,35]

Increasingly, the quality of evidence is assessed by GRADE (Grading of Recommendations, Assessment, Development and Evaluations). GRADE, a systematic and explicit approach to forming judgments about the quality of evidence and strength of recommendations, is widely seen as an valuable method of linking evidence-quality evaluations to clinical recommendations.[36] Treatment comparisons are rated on GRADE scores that reflect the quality of the evidence. The scores are also included in guideline recommendations. This provides information on strong and weak recommendations that might help patients and healthcare providers in decision making, and is therefore valuable in guiding clinical practice. The AGREE instrument (Appraisal of Guidelines in Research and Evaluation) may facilitate the involvement of patients in guideline development.[37,38] The AGREE II Instrument is an appraisal tool for evaluating guideline development and reporting. With respect to the appraisal tools that cover several different aspects of guideline quality, rigorous development does not guarantee appropriate recommendations or better health outcomes for patients, as the methodological rigor and quality of the contents of a clinical practice guideline are not necessarily correlated. [39,40] Even if a guideline's recommendations are perfectly concordant with the best evidence available, providers may still be

dissatisfied if the recommendations are hard to apply, or they may leave important clinical problems unaddressed because of gaps in the literature.

Guideline developers have to keep in mind that healthcare providers, as the primary end-users of guidelines, prefer guidelines that are applicable and practical. The guideline development procedure must, therefore, focus on acceptance by the guideline users. Guidelines should encourage clinical reasoning and awareness by: 1) supportive questions that facilitate hypothesis testing and reasoning; 2) integration of an ICF framework to identify all interacting factors that are related to the patients' health; 3) the inclusion of PROMs and other outcome measures with time frames for monitoring the course and evaluation of hypotheses, goals and treatment plans.

Electronic patient records can support the physical therapist in their record keeping, as the use of electronic patient records can improve the quality of healthcare by increasing time efficiency and guideline adherence and reducing errors.[41] The inclusion of guidelines with clinical decision support systems in the electronic patient records serves to help physical therapists apply the guidelines by providing up-to-date medical knowledge, reminders or other actions that aid health professionals in decision making.[41] Those decision tools have become increasingly sophisticated by matching patient characteristics with computerized knowledge bases and using algorithms to generate patient-specific assessments or treatment recommendations.[42] Patients can also have a role in reporting their health status and progress. Patient information on the history of complaints and PROMs can be completed online beforehand by patients, and reminders can be sent automatically.[43] This will save time, improve the completeness of patient files, and encourage reflection on the outcomes.

### *2. Implications for professionals*

Guidelines help physical therapists in recognizing typical and atypical signs and symptoms. That, in turn, will help them understand causal relationships between ICF levels, gain insight into the expected course of recovery and weigh pros and cons in decision making. Physical therapists need to develop a routine of self-questioning and hypotheses evaluation during their clinical work. By promoting awareness, reflection and critical appraisal, clinical reasoning can be enhanced, and physical therapists will become interested in developing their own approach to making diagnoses and decisions.[15]

Physical therapists are supposed to adopt an evidence-based attitude towards medical education and a continuous professional development process. With the proper specialized training, physical therapists can learn to detect and correct their own reasoning flaws. The continuing professional development process may include learning portfolios, documenting practice-based learning and improving activities, working with multisource feedback, engaging in interactive self-assessment focused

on medical knowledge, and evaluating performance in practice. Physical therapists are expected to learn to reason aloud with peers or during formal mentoring opportunities, using hypotheses and monitoring expectations, and to practice their profession based on a model of critical thinking and of receiving and providing feedback.

It is envisioned that PROMs will help improve the quality of care, as PROMs can assist physical therapists in different areas, including: 1) their clinical reasoning during the intake process; 2) their identification of needs; 3) their diagnoses; 4) their process of gaining insight into signs and symptoms and 5) their treatment and monitoring of outcomes, a task focused on the patient's perspective.[24] Physical therapists should use PROMS when recommended in guidelines. The PROM should be well coordinated with the patient's need for help (e.g. pain activity) and the treatment goals. Generic PROMs, such as the PSC questionnaire, or the Numeric Pain Rating Scale (NPRS) are feasible in the majority of patient groups that visit primary care physical therapy.

Another factor is the role patient records may have in the collection and optimal use of data. The sharing of information between patients and physical therapists will require further technological development. Active involvement on the part of patients, physical therapists and the Royal Dutch Society of Physical Therapy is essential in this process. The national program for quality improvement that was initiated by the KNGF should make this element a high priority (see Box 3).

### *3. Implications for the patient*

A new concept of health was introduced in 2011 by Huber et al.: "Health as the ability to adapt and to self-manage, in the face of social, physical and emotional challenges".[44] This new concept was proposed because the old, traditional WHO definition of health was no longer considered adequate. Formulated in 1948, that definition reads: "Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity".[45] Clearly, this statement described a static state, according to which almost everybody could be considered ill, to some extent; as such, this definition unintentionally enhanced medicalization.[44] The new definition puts forward an increasing responsibility for the patient's own health, one that requires an active role on the part of the patient and health professionals. To tailor goals and meet the intervention needs and preferences of patients, as well as to optimize patient-centered care, several actions can be taken. First, patient participation in clinical practice can be improved by enhancing patient education and physical therapist training. This will facilitate a change in the attitude of patients and physical therapists towards sharing authority and responsibility.[46] Secondly, patient versions of guidelines can support patients as they are aimed at involving them in the process of healthcare improvement, either by learning about the current standard of care, or

**Box 3** Illustration of the Dutch physical therapy quality improvement program

In 2013, the Royal Dutch Society for Physical Therapy (KNGF) initiated a national program for quality improvement. The program's introduction was prompted by the desire among this professional body to establish a quality assurance program that would stimulate continual professional development by ensuring transparency in care based on health outcomes measurements. The program's purpose is multifold:

- To implement the quality assurance program, a strategy based on peer assessment was developed in order to improve clinical reasoning strategies and the use of health outcomes measurement that support physical therapists and patients in goal setting, shared decision making and monitoring of the patient's health throughout the care process, thus enhancing patient centeredness.
- Patient data based on clinical guidelines is collected electronically a part of normal daily practice in electronic patient files. This routine data collection will enable ongoing evaluation of current care, and will provide valuable information that can be used to implement and update guidelines.
- Data collection at the national level can be used for quality improvement activities, for the benchmarking process and for overseeing the outcomes of physical therapy care.
- To allow for reflections on the outcomes of care and to develop improvement strategies, the benchmarking results of individual therapists or practices are presented in an online feedback tool.

by enabling them to make informed decisions on their health, supported by the best evidence available. Thirdly, individualized measures enhance patient participation, as they can facilitate more dynamic interaction between the patient and the physical therapist, allowing the identification of personal values and priorities.[47] For comparing the results of the intervention, a post-treatment measurement is of great importance. This allows the physical therapist to measure the effect of the treatment, and to compare the results between patients at the aggregate level. This is becoming more important, as there is a societal demand for accountability in healthcare and transparency as regarding the quality of the care delivered. Patients should be aware of this and complete the post-treatment measurements, although this may not be used directly in evaluating their treatment, e.g., when the post-treatment measurement is completed after the final treatment session.

### *Implications for research*

Given the study findings in this thesis, we would recommend the following directions for future research:

- Future research is needed to identify the possibilities of electronic patient records support in the clinical reasoning process and in patient involvement, and to facilitate and optimize record keeping, as well as to minimize the burden of registration. One question that needs to be addressed is: how can guideline recommendations be incorporated in the electronic patient records to ensure that it facilitates diagnostics, goal setting, evaluation and reflection. New developments, such as spoken language technology, should be tested.
- Research is required to identify how information about the process and outcome of care can be used for quality improvement. Technological developments make it possible to build large databases that can identify the following: 1) the data needed to provide insight into effective interventions and outcomes; 2) clinically important improvements on PROMs for various patient groups and 3) how these measures can be used to create an environment for reflection with peers.
- While the use of PROMs for different purposes is potentially meaningful and supportive, additional research is necessary to identify PROMs that: 1) enhance patient participation, shared decision-making, and treatment evaluation at the individual level and 2) contribute to the reporting of treatment effectiveness for quality improvement purposes at the group level. Exploration of the information needs, barriers to participation, incentives, expectations and experiences of healthcare providers and patients, will provide important information to support the future application of PROMs. Focused education, for both students and practicing professionals, are necessary to establish the standard use of outcome measures in practice.
- In order to improve the quality of healthcare based on the findings of our study, a strategy should be developed and tested for physical therapists for the implementation of peer assessment. Expert assessors should be trained and elements of the successful peer assessment strategy should be complemented with other components of quality in care, such as communication with the patient, record keeping, and the use of PROMs in clinical decision making.
- To enhance patient-centered care, additional research is required on how physical therapists and patients can change their attitudes towards sharing responsibility in goal setting and treatment decisions.

## Conclusion

To improve patient-centered quality in healthcare, new initiatives and formats are required to establish a culture of reflective practice and lifelong learning. To ensure that data in patient records lead to insight into the physical therapist's behavior, policy makers must pay more attention to the delivery of feedback and organizational support in data collection. Moreover, guideline developers should include tools to support clinical reasoning and routine outcome measurements in using clinical practice guidelines. By promoting awareness, reflection and critical appraisal, learning needs will be identified. Quality improvement is a process that will require physical therapists to adopt an attitude of systematically examining their practice performance in order to identify clinical reasoning skills and knowledge gaps.

Physical therapists are expected to learn to reason aloud with peers or during formal mentoring opportunities, using hypotheses and monitoring expectations, and to practice their profession based on a model of critical thinking and of receiving and providing feedback.

# References


1. van Dulmen SA, Lukersmith S, Muxlow J, Santa Mina E, Nijhuis-van der Sanden MW, van der Wees PJ, on behalf of the GIN Allied Health Steering Group. Supporting a person-centred approach in clinical guidelines. A position paper of the Allied Health Community - Guidelines International Network (G-I-N). Health Expect 2015; 18(5): 1543-58.
2. van Dulmen SA, Tacken MA, Staal JB, Gaal S, Wensing M, Nijhuis-van der Sanden MW. Patient safety in primary allied health care: what can we learn from incidents in a Dutch exploratory cohort study? Med Care 2011; 49(12): 1089-96.
3. van Dulmen SA, Maas M, Staal JB, Rutten G, Kiers H, Nijhuis-van der Sanden M, van der Wees P. Effectiveness of peer assessment for implementing a Dutch physical therapy low back pain guideline: cluster randomized controlled trial. Phys Ther 2014; 94(10): 1396-409.
4. Maas MMJ, Van Dulmen SA, Sagasser MH, Heerkens YF, Van der Vleuten CMP, Nijhuis-van der Sanden MWG, Van der Wees PJ. Critical features of peer assessment of clinical performance to enhance adherence to a low back pain guideline for physical therapists: a mixed methods design. BMC Med Educ 2015; 15(1): 203.
5. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. J Gen Intern Med 2003; 18(1): 61-7.
6. de Bont AA, Jerak S, de Mul M, Zwart D. Vragen voor veiligheid. De betekenis van patiëntveiligheid in de eerste lijn. [Questions for safety. The meaning of patient safety in primary care]. In, Rotterdam: Erasmus University Rotterdam, 2009.
7. Davis DA, Mazmanian PE, Fordis M, Van Harrison R et al. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA 2006; 296(9): 1094-102.
8. Maas MJ, van der Wees PJ, Braam C, Koetsenruijter J, Heerkens YF, van der Vleuten CP, Nijhuis-van der Sanden MW. An innovative peer assessment approach to enhance guideline adherence in physical therapy: single-masked, cluster-randomized controlled trial. Phys Ther 2015; 95(4): 600-12.
9. Del Mar C, Doust J, Glasziou P. Clinical Thinking: Evidence, Communication and Decision-Making. Hoboken, United States: Blackwell Publishing Ltd, 2006.
10. Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. Am J Med 2008; 121(5 Suppl): S2-23.
11. Nendaz M, Perrier A. Diagnostic errors and flaws in clinical reasoning: mechanisms and prevention in practice. Swiss Med Wkly 2012; 142w13706.
12. Edwards I, Jones M, Carr J, Braunack-Mayer A, Jensen GM. Clinical reasoning strategies in physical therapy. Phys Ther 2004; 84(4): 331-15.
13. Zwart DLM, De Bont AA. Introducing incident reporting in primary care: a translation from safety science into medical practice. Health, Risk & Society 2013; 15(3): 265- 78
14. Snijders C, Kollen BJ, van Lingen RA, Fetter WP et al. Which aspects of safety culture predict incident reporting behavior in neonatal intensive care units? A multilevel analysis. Crit Care Med 2009; 37(1): 61-7.
15. Higgs J, Jones MA, Loftus S, Christensen N. Clinical reasoning in the health professions. Oxford: Butterwordth Heineman, 2008.
16. Wainwright SF, Shepard KF, Harman LB, Stephens J. Novice and experienced physical therapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. Phys Ther 2010; 90(1): 75-88.
17. Escorpizo R, Bemis-Dougherty A. Introduction to Special Issue: A Review of the International Classification of Functioning, Disability and Health and Physical Therapy over the Years. Physiother Res Int 2015 Dec;20(4):200-9.
18. Cieza A, Stucki G, Weigl M, Disler P et al. ICF Core Sets for low back pain. J Rehab Med 2004(44 Suppl): 69-74.
19. Stucki G, Kostanjsek N, Ustun B, Cieza A. ICF-based classification and measurement of functioning. Eur J Phys Rehabil Med 2008; 44(3): 315-28.
20. Spoorenberg SL, Reijneveld SA, Middel B, Uittenbroek RJ et al. The Geriatric ICF Core Set reflecting health-related problems in community-living older adults aged 75 years and older without dementia: development and validation. Disabil Rehabil 20151-7.

21. Yen TH, Liou TH, Chang KH, Wu NN et al. Systematic review of ICF core set from 2001 to 2012. Disabil Rehabil 2014; 36(3): 177-84.
22. Rothstein JM, Echternach JL, Riddle DL. The Hypothesis-Oriented Algorithm for Clinicians II (HOAC II): a guide for patient management. Phys Ther 2003; 83(5): 455-70.
23. Resnik L, Hart DL. Using clinical outcomes to identify expert physical therapists. Phys Ther 2003; 83(11): 990-1002.
24. Kyte DG, Calvert M, van der Wees PJ, Ten Hove R et al. An introduction to patient-reported outcome measures (PROMs) in physiotherapy. Physiotherapy 2014; 101(2): 119-25.
25. Greenhalgh J, Meadows K. The effectiveness of the use of patient-based measures of health in routine practice in improving the process and outcomes of patient care: a literature review. J Eval Clin Pract 1999; 5(4): 401-16.
26. Furlan AD, Giraldo M, Baskwill A, Irvin E, Imamura M. Massage for low-back pain. Cochrane Database Syst Rev 2015; 9CD001929.
27. Campbell E, Coulter EH, Mattison PG, Miller L et al. Physiotherapy rehabilitation for people with progressive Multiple Sclerosis: a systematic review. Arch Phys Med Rehabil 2016; Jan;97(1):141-51.
28. Black N. Patient reported outcome measures could help transform healthcare. BMJ 2013; 346:f167.
29. Boivin A, Currie K, Fervers B, Gracia J et al. J. Patient and public involvement in clinical guidelines: international experiences and future perspectives. Qual Saf Health Care 2010; 19(5): e22.
30. Burgers JS, Grol R, Klazinga NS, Makela M et al. Collaboration A. Towards evidence-based clinical practice: an international survey of 18 clinical guideline programs. Int J Qual Health Care 2003; 15(1): 31-45.
31. Chong CA, Chen IJ, Naglie G, Krahn MD. How well do guidelines incorporate evidence on patient preferences? J Gen Intern Med 2009; 24(8): 977-82.
32. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. Health Res Policy Syst 2006; 422.
33. Blauwdruk patiëntenparticipatie in richtlijnontwikkeling. Leidraad voor richtlijnmakers. In, 2010.
34. Hilbrink M, Ouwens M, kool T. De haring-tools. Diertien instrumetne voor onderstuening bij het opstellen, herzien, implementeren en evalueren van richtlijnen. Nijmegen: IQ healthcare, Radboudumc, 2013.
35. Hilbink MA, Ouwens MM, Burgers JS, Kool RB. A new impetus for guideline development and implementation: construction and evaluation of a toolbox. Implement Sci 2014; 9(1): 34.
36. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 2008; 336(7650): 924-6.
37. Brouwers MC, Kho ME, Browman GP, Burgers JS et al. Development of the AGREE II, part 1: performance, usefulness and areas for improvement. Can Med Assoc J 2010; 182(10): 1045-52.
38. Collaboration A. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003; 12(1): 18-23.
39. Watine J, Friedberg B, Nagy E, Onody R et al. Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. Clin Chem 2006; 52(1): 65-72.
40. Nuckols TK, Lim YW, Wynn BO, et al. Rigorous development does not ensure that guidelines are acceptable to a panel of knowledgeable providers. J Gen Intern Med 2008; 23(1): 37-44.
41. Campanella P, Lovato E, Marone C, Fallacara L et al. The impact of electronic health records on healthcare quality: a systematic review and meta-analysis. Eur J Public Health 2015.
42. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA 2005; 293(10): 1223-38.
43. Hayrinen K, Saranto K, Nykanen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. Int J Med Inform 2008; 77(5): 291-304.
44. Huber M, Knottnerus JA, Green L, van der Horst H et al. How should we define health? BMJ 2011; 343d4163.
45. WHO. Constitution of the World Health Organization. In, 2006.
46. Schoeb V, Burge E. Perceptions of patients and physiotherapists on patient participation: a narrative synthesis of qualitative studies. Physiother Res Int 2012; 17(2): 80-91.
47. Haywood K, Marshall S, Fitzpatrick R. Patient participation in the consultation process: a structured review of intervention strategies. Patient Educ Couns 2006; 63(1-2): 12-23.

# 8

# Summary

**Chapter 1** presents an overview of this thesis, describing the factors important to improve the quality of primary physical therapy care. It addresses important elements of quality care in general, and more specifically, in physical therapy care. This thesis explores three dimensions of quality in healthcare: safety, effectiveness and patient centeredness. *Safety* means that patients should not be harmed by the care that is intended to help them. *Effectiveness* is associated with evidence-based practice in which clinical expertise, patient values, and the best available evidence is integrated into the decision-making process for patient care. By translating the best available evidence into specific guideline recommendations, guidelines can facilitate the uptake of new research findings and insights into clinical practice, thereby facilitating effective care. *Patient centeredness* is defined as providing care that is responsive to individual patient preferences, needs and values, and where the patient plays an active role in making decisions about his/her own care.

Multiple elements at different levels of healthcare (e.g. organizational, policy, healthcare provider, patient) are relevant to gaining insight into the quality of care. These insights will guide strategies for improving the quality of healthcare. The focal research question in this thesis is: what factors can be influenced to improve the quality of healthcare in Dutch physical therapy practice ?

**Chapter 2** focuses on identifying patient safety incidents in primary allied healthcare in the Netherlands, as there is hardly any concrete information on patient safety in primary allied healthcare. In a preliminary study, a questionnaire was developed to test the feasibility of the patient record study and to trace the possible nature of incidents. The results of this questionnaire showed that allied health care therapists (occupational therapists, physical therapists and Cesar/Mensendieck exercise therapists linked situations that were (potentially) unsafe to patients mainly to patient-related aspects, followed by a lack of safe equipment and an inadequate history. Subsequently, a retrospective study of 1000 patient records in a representative sample of 20 allied healthcare practices was combined with a prospective incident-reporting study. All records were reviewed by trained researchers to identify patient safety incidents. In reviewing the incidents, we applied the Prevention and Recovery Information System for Monitoring and Analysis (PRISMA) method to analyze them by means of causal factor trees. This taxonomy distinguishes 5 main categories of causes: technical, organizational, human, patient-related, and other, which can be subdivided into subcategories. We identified 18 incidents out of 1000 patient records in 11 physical therapist practices (550 records), 6 exercise therapy practices (300 records), and 3 occupational therapy practices (150 records). The main causes of incidents were related to errors in clinical decisions (89%), communication with other healthcare providers (67%), and monitoring (56%). The probability of incidents was higher if more healthcare providers had been involved, and if patient records

were incomplete (37% of the records). No incidents were reported in the prospective study. We concluded that although the absolute number of incidents was low, the incompleteness of the patient records may have resulted in an underestimation of incidents. Therapists were unaware of incidents and causes, and consequently had not personally observed or reported unsafe situations. Many patients were not treated according to the guidelines, and intervention results were not adequately monitored and evaluated. The fact that the incidents were mainly attributable to human actions suggests that a focus on clinical reasoning and record keeping is needed to further enhance patient safety.

**Chapter 3** presents the results of our cluster-randomized controlled trial testing of the effectiveness of peer assessment for implementing a Dutch physical therapy guideline for managing low back pain in patients. Peer assessment is a strategy in which professionals assess (judge) the performance of their peers, using relevant criteria and providing structured feedback in a simulated setting. Our aim was to determine whether peer assessment is more effective than case-based discussions to improve knowledge and guideline-consistent clinical reasoning. The main difference between peer assessment and case-based discussions is that peer assessment intervention focuses on the assessment of performance rather than on discussions. Ten communities of practice (CoPs) of physical therapists (N=90) were cluster randomized: 6 CoPs in the peer assessment group (n=49) and 4 CoPs in the case-based discussion group (control group) (n=41). Both groups participated in 4 educational sessions and used clinical patient cases. The peer assessment group reflected on performance in low back pain management among parties with different roles: patients, physical therapists, and assessors. Performance was assessed with a scoring sheet containing performance criteria evaluated on a 7-point scale and some space for qualitative feedback. The primary outcome measure was knowledge and guideline-consistent reasoning, measured with performance indicators using 4 clinical vignettes with specific guideline-related patient profiles. Per vignette and for each step in clinical decision making, a performance indicator was used to measure guideline knowledge and guideline-consistent clinical reasoning, for a total of 12 indicators. The secondary outcome measure was reflective practice, as measured by the Self-Reflection and Insight Scale. Multi-level analysis showed an increase in guideline-consistent clinical reasoning of 8.4% in the peer assessment group, whereas the control group showed a decline of 0.1% (estimated group difference = 8.7%, 95% confidence interval = 3.9 to 13.4). No group differences were found on self-reflection. In conclusion, this trial demonstrated that peer assessment leads to an increase in knowledge and guideline consistent clinical reasoning. This effect may be explained by the combination of different educational strategies: dissemination of the guideline, in-depth assessment of the guideline in a problem-solving process,

assessment of performance in different roles, individualized well-timed performance feedback, and an individually tailored improvement plan.

Based on the results of the cluster-randomized trial, an in-depth evaluation was performed to identify the critical features of the peer assessment strategy that were perceived to have a powerful impact on learning and change of routine practice in **Chapter 4**. We conducted a mixed methods study with physical therapists that participated in the peer-assessment groups in the trial. We used questionnaires and semi-structured interviews to explore the attributes of peer-assessment that contributed to improved guideline adherence in the perception of participants.

The peer-assessment program was decomposed into learning tasks and subtasks. After the program was finished, a questionnaire was administered in which participants were asked to rank the subtasks from high to low learning value and to motivate their choices. Additional semi-structured interviews were conducted to elaborate on the questionnaire results. The results showed that subtasks related to performance in the therapist role were perceived to have the highest impact on learning, although task perceptions varied from challenging to threatening. Although some participants initially felt reluctant to move out of their "comfort zone", they considered exposure of their routine practice as a necessity for quality improvement. In the PT role, participants needed to make the transfer from implicit reasoning to explicit reasoning and from intentional behavior to observable behavior to allow for assessment and feedback. Participants showed a strong cognitive and emotional commitment to performing the tasks related to the physical therapist role. Learning outcomes were awareness of strengths and weaknesses in clinical performance, improved attitudes towards the guideline, and increased self-efficacy beliefs in managing patients with low back pain. Task perceptions were affected by the role-play format and the time schedule. Learning was facilitated by psychological safety and the quality of feedback. A safe learning environment was facilitated by the coach by means of posing critical questions rather than giving straightforward answers. We concluded that the critical success factor of peer-assessment can be attributed to the performance based-design provoking a strong cognitive and emotional involvement with the therapist role and the assessor role.

**Chapter 5** describes the development of a position paper of the Allied Health Community of the Guidelines International Network (G-I-N). G-I-N is a global network that supports evidence- based healthcare and improvement of health outcomes by promoting the development, implementation and use of clinical practice guidelines. One of the aims of the G-I-N Allied Health Community is to promote patient-centered health services and to promote health-related quality of life activities in clinical guidelines. The aim of this study is to develop a position paper that promotes a

person-centered approach in guideline development and implementation. We used three narrative discussion formats to collect data for achieving consensus: a nominal group technique for the G-I-N Allied Health Steering Group, an Internet discussion board and a workshop at the annual G-I-N conference. We analyzed the data for relevant themes to draft recommendations. We built the position paper on the values of the biopsychosocial model. Four key themes for enhancing a person-centered approach in clinical guidelines emerged: (i) use a joint definition of health-related quality of life as an essential component of intervention goals, (ii) incorporate the International Classification of Functioning, Disability and Health (ICF) as a framework for considering all domains related to health, (iii) adopt a shared decision-making method, and (iv) incorporate patient-reported health outcome measures. Person-centered care focuses sharply on participation of the person in clinical decision making by taking into account his/her perspective and by tailoring health services to the needs and preferences of the person. The consensus process provided information about barriers and facilitators that might help us to develop strategies for implementing person-centered care. The position statement includes 14 recommendations for guideline developers, implementers, and users.

**Chapter 6** reports the findings of an observational cohort study in using patient-reported outcome measures (PROMs) for goal setting and outcome measurement in physical therapy practices. The routine use of PROMs is of increasing interest, because PROMs might facilitate goal setting with the patient in selecting health outcomes of the highest priority, and PROMs may provide an effective way of monitoring patient-valued outcomes. In this study we explored 1) the current use of PROMs in patients with neck and low back pain; 2) to what extent the goals correspond with the selected PROMs; 3) the health outcomes based on PROM measurements. In total, 43 practices with 60 physical therapists participated. All information regarding the treatment episode of each patient was recorded in an EHR. The PROMs selected were the Visual Analogue Scale (VAS) for pain measures. For limitations in activities and participations the Patient Specific Complaints (PSC), (the Quebec Back Pain Disability Scale (QBPDS), and the Neck Disability Index (NDI) were used. The physical therapists identified goals based on the history taking (including the scores on the PROMs) and diagnoses. These goals were recorded on a pre-structured form based on a shortlist of categories of the International Classification of Functioning, Disability and Health (ICF). We analyzed data from 299 patients with neck pain or low back pain. One baseline measurement was completed by 70% of the patients with neck pain, and a repeated measurement was completed by 61% of the patients. One baseline measurement was completed by 68% of the patients with low back pain, and a repeated measurement was completed by 63% of the patients. In 46% of the patients with a pain reduction goal, a VAS was used, whereas in 43% of the patients

with an activity or participation goal, a QBPDS, NDI or PSC was used. The high number of treatment goals (median 7 per patient) explained by the overlap between the different items, which makes it difficult to value the most important goals and to assess which PROM is most appropriate. The mean differences between baseline and follow-up scores for all PROMs were statistically significant, and the majority of the patients showed improved health outcomes. We concluded that the results showed that the PROMs were used in the majority of the patients and corresponded moderately with the goals. Further research is necessary to identify how PROMs are used in the goal setting process.

**Chapter 7**, the final chapter of this thesis, discusses the most important findings and conclusions of the studies. The results are placed in wider perspectives and compared with other studies. The most relevant implications for guideline development and implementation are considered, implications for the physical therapists and patients are discussed, and recommendations for future research are outlined.
Improvement of the quality and safety of the care delivered requires physical therapists to systematically reflect on their practice performance in order to identify gaps in knowledge and clinical reasoning skills. Physical therapists need to develop an attitude of critical reflection, and establish a routine of self-questioning during clinical work. By promoting awareness and reflection, clinical reasoning can be enhanced. Physical therapists should adopt a positive attitude towards continuous professional development, and in using guidelines. To ensure that data in patient records lead to insight in the physical therapist's behavior, policy makers must pay more attention to the delivery of feedback and organizational support in data collection. Guideline developers should include tools to support clinical reasoning and outcome measurements with PROMs with time frames for monitoring the course and re-analysis of the hypothesis, goals and treatment plans. Patient versions of guidelines will enhance patient involvement. Guideline implementation is crucial for facilitating evidence-based practice.

# 9

Nederlandse samenvatting

Dankwoord

Over de auteur

PhD portfolio

List of publications

# Nederlandse samenvatting

Er bestaat veel aandacht voor de kwaliteit van de zorg en verschillende partijen hebben ieder een specifieke invalshoek op kwaliteit. Zo kunnen patiënten een andere visie op kwaliteit hebben dan zorgverleners of zorgverzekeraars. Dit proefschrift beschrijft de verschillende elementen van kwaliteit van de eerstelijns fysiotherapeutische zorg en factoren die de kwaliteit kan verbeteren.

**Hoofdstuk 1** beschrijft belangrijke elementen van de kwaliteit van de zorg in het algemeen, en in het bijzonder van de fysiotherapeutische zorg. Kwaliteit van zorg kan verschillende elementen bevatten, zoals veiligheid, effectiviteit, patiëntgerichtheid, efficiëntie, tijdigheid en gelijkheid. Dit proefschrift onderzoekt de drie eerstgenoemde elementen.

Veiligheid betekent het vermijden van schade door zorg die is bedoeld om patiënten te helpen. Effectiviteit betreft het leveren van zorg op basis van de meest recente wetenschappelijke kennis, klinische expertise en voorkeuren van de patiënt. Patiëntgerichtheid wordt gedefinieerd als de zorg die rekening houdt met de individuele voorkeuren van de patiënt, de behoeften en waarden, en waarbij de patiënt een actieve rol speelt bij het nemen van beslissingen over zijn / haar eigen zorg.

Een belangrijk instrument voor het overbruggen van de kloof tussen wetenschappelijk bewijs en de klinische praktijk, zijn richtlijnen. Deze zijn bedoeld om patiëntgerichte zorg te ondersteunen en verbeteren. Door het vertalen van wetenschappelijk bewijs in specifieke aanbevelingen in richtlijnen, worden zorgverleners ondersteund in de dagelijkse praktijkvoering.

Verschillende stakeholders (bijvoorbeeld organisatie, zorgverlener en patiënt) kijken op specifieke manieren aan tegen veiligheid, effectiviteit en patiëntgerichtheid. Inzicht in de kwaliteitsaspecten is nodig om strategieën te ontwikkelen om de kwaliteit van de fysiotherapeutische zorg te verbeteren.

De centrale onderzoeksvraag in dit proefschrift is: 'Welke factoren kunnen worden beïnvloed om de kwaliteit van de eerstelijns fysiotherapie te verbeteren?'

In **hoofdstuk 2** is de patiëntveiligheid binnen drie paramedische beroepsgroepen in kaart gebracht. Hoewel er in Nederland veel zorg in de eerste lijn plaatsvindt, is er weinig bekend over de patiëntveiligheid in de eerstelijnsgezondheidszorg. In een eerste studie is een vragenlijst ontwikkeld om de haalbaarheid van het dossieronderzoek te testen en om de aard van de incidenten op te sporen. Uit de resultaten bleek dat de paramedische zorgverleners (ergotherapeuten, fysiotherapeuten en Cesar / Mensendieck oefentherapeuten) potentieel onveilige situaties hoofdzakelijk koppelden aan patiëntgerelateerde aspecten, een gebrek aan veilige apparatuur en een onvolledige anamnese. De methode van dossieronderzoek werd zowel op inhoud als op

praktische uitvoerbaarheid getest in drie pilots bij paramedische praktijken. Vervolgens is retrospectief onderzoek gedaan op basis van 1000 patiëntendossiers uit 20 paramedische praktijken. Daarnaast konden zorgverleners 2 weken (prospectief) incidenten melden. Een incident is gedefinieerd als 'een onbedoelde gebeurtenis tijdens het zorgproces die tot schade aan de patiënt heeft geleid, had kunnen leiden of (nog) kan leiden'. De gevonden en gemelde incidenten werden op een methodische wijze door de onderzoekers geanalyseerd naar typeringen, oorzaken en gevolgen. De oorzaken van de incidenten werden ingedeeld volgens het Eindhovens Classificatie Model (onderdeel van de PRISMA-methode). We identificeerden 18 incidenten uit 1000 patiëntendossiers in 11 fysiotherapiepraktijken (550 dossiers), 6 oefentherapiepraktijken (300 dossiers), en 3 ergotherapie praktijken (150 dossiers). De belangrijkste oorzaken van incidenten hadden betrekking op fouten in de klinische beslissingen (89%), de communicatie met andere zorgverleners (67%), en monitoring (56%). De kans op incidenten was hoger als er meer zorgverleners betrokken waren en als patiëntendossiers onvolledig waren. In de prospectieve studie werden geen incidenten gerapporteerd.

Het absolute aantal incidenten dat naar voren kwam uit de dossiers, was laag. Desondanks kon worden geconcludeerd dat onvolledigheid van de patiëntendossiers kan hebben geleid tot een onderschatting van incidenten. Zorgverleners waren niet zich niet bewust van onveilige situaties en de oorzaken van de incidenten. Veel patiënten waren niet behandeld volgens richtlijnen, en het beloop van de klachten werd niet adequaat gecontroleerd en geëvalueerd. Het feit dat de incidenten vooral te wijten waren aan menselijk handelen, suggereert dat een focus op klinisch redeneren en registratie nodig is om de patiëntveiligheid te verbeteren.

**Hoofdstuk 3** beschrijft een cluster gerandomiseerde trial waarbij is gekeken of peer assessment effectiever is dan casusbesprekingen om kennis en klinisch redeneren te verbeteren. Peer assessment is een strategie waarbij zorgverleners de prestaties van hun collega's (peers) beoordelen. Zij geven in een gesimuleerde omgeving feedback aan de hand van criteria. In tegenstelling tot casusbesprekingen, richt de peer assessment-interventie zich op het geven van gestructureerde feedback en de beoordeling van de prestaties in plaats van op discussies. Bij de studie werd de effectiviteit van peer assessment onderzocht bij de implementatie van de (herziene) Nederlandse fysiotherapierichtlijn voor patiënten met lage rugpijn.

Tien IOF's (Intercollegiaal Overleg Fysiotherapie) van fysiotherapeuten (n = 90) werden in clusters gerandomiseerd: 6 IOF's in de peer assessment-groep (n = 49) en 4 IOF's in de casusbesprekinggroep (controlegroep) (n = 41). Beide groepen namen deel aan vier educatieve sessies en gebruikten patiëntencasussen. De zorgverleners uit de peer assessment-groep beoordeelden de prestaties van hun collega's vanuit verschillende rollen: als patiënt, als fysiotherapeut en als beoordelaar.

Prestaties werden ingevuld op een beoordelingsformulier. De primaire uitkomstmaat was de toename van kennis en richtlijnconsistent redeneren, gemeten met indicatoren op basis van 4 klinische vignetten met specifieke patiëntprofielen. Per vignet en voor elke stap in de klinische besluitvorming is een prestatie-indicator gebruikt om de richtlijn kennis en klinisch redeneren te meten, in totaal 12 indicatoren. De secundaire uitkomstmaat was reflectie, gemeten met de Self-Reflection and Insight Scale (SRIS). De SRIS is een instrument om het proces van zelfreflectie en inzicht te meten, wat wordt gezien als voorwaardelijk voor zelfgerichte veranderingen. De gemiddelde score op de voormeting bij de peer assessment-groep was 63,7% en bij de nameting 72,0%. De controlegroep had een gemiddelde voormeting van 66,8% en een nameting van 66,7%. Multilevel analyse toonde bij de peer assessment-groep een toename van de kennis en klinisch redeneren aan van 8,4% en bij de controlegroep een daling van 0,1% (geschatte groepsverschil = 8,7%, 95% betrouwbaarheidsinterval = 3,9-13,4). We vonden geen verschillen in zelfreflectie.

We concludeerden dat peer assessment leidt tot een toename van kennis en het volgen van de richtlijn. Dit effect kan worden verklaard door de combinatie van verschillende educatieve strategieën: de verspreiding van de richtlijn, de diepgaande evaluatie van de richtlijn, het handelen in een gesimuleerde setting, de beoordeling van de prestaties in verschillende rollen, geïndividualiseerde feedback, en een individueel verbeterplan.

**Hoofdstuk 4** beschrijft de diepteanalyse van resultaten van de cluster gerandomiseerde trial uit hoofdstuk 3. Het doel was de kritische succesfactoren van de peer assessment-strategie te identificeren en zo inzichtelijk te maken welke kenmerken van de peer assessment-strategie hebben bijgedragen aan de toename van de kennis en het volgen van de richtlijn. Voor deze diepteanalyse is eerst onderscheid gemaakt tussen taken die horen bij de verschillende rollen (fysiotherapeut, patiënt of assessor) van het peer assessment-programma. De fysiotherapeuten die hadden deelgenomen aan dit programma ontvingen een vragenlijst waarin ze een score konden toekennen aan de taken en hun keuze konden motiveren. Daarnaast werden semigestructureerde interviews gehouden om de uitkomsten van de vragenlijst verder te duiden.

De resultaten toonden aan dat de rol als fysiotherapeut de grootste invloed op leren had en daarmee de belangrijkste succesfactor was een toename van kennis. Het ontvangen van peer feedback was het meest gewaardeerde element, gevolgd door de feedback van de externe coach en het uitvoeren van de rol als fysiotherapeut in het rollenspel. Hoewel sommige deelnemers in eerste instantie terughoudend waren, zagen ze het praktisch uitvoeren van hun rol en het hardop klinisch redeneren als belangrijke elementen voor kwaliteitsverbetering. De deelnemers toonden een sterke cognitieve en emotionele betrokkenheid bij het uitvoeren van de taken in de rol als fysiotherapeut. Voor de fysiotherapeut waren de belangrijkste leeraspecten van

de peer assessment-strategie de bewustwording van de sterke en zwakke punten in het klinisch redeneren, een betere houding ten opzichte van de richtlijn, en de toegenomen bewustwording van de beperkingen van kennis en vaardigheden. Het leren werd beïnvloed door psychologische veiligheid en kwaliteit van feedback. Ten slotte werd de gestructureerde opzet van de peer assessment-strategie als een belangrijke succesfactor gezien.

**Hoofdstuk 5** beschrijft de ontwikkeling van een position paper waarmee een persoonsgerichte aanpak in richtlijnontwikkeling en implementatie wordt bevorderd. Dit paper werd ontwikkeld door de Allied Health Community van het Guidelines International Network (G-I-N). G-I-N is een wereldwijd netwerk dat evidence-based gezondheidszorg en de verbetering van de gezondheidsuitkomsten ondersteunt door het bevorderen van de ontwikkeling, de implementatie en het gebruik van de klinische richtlijnen. Persoonsgerichte zorg richt zich sterk op de deelname van de patiënt in de klinische besluitvorming door rekening te houden met zijn / haar perspectief en door het afstemmen van de zorg aan de behoeften en voorkeuren van de patiënt.

Om data te verzamelen werd gebruik gemaakt van drie narratieve discussie-methoden: een nominale groepstechniek met de G-I-N Allied Health Steering Group, een internet-discussieforum en een workshop tijdens de jaarlijkse G-I-N conferentie. We analyseerden de gegevens op relevante thema's om aanbevelingen te formuleren. Deze waren: (1) zet gezondheidsgerelateerde kwaliteit van leven in als een essentieel onderdeel van de interventiedoelen, (2) gebruik de International Classification of Functioning, Disability and Health (ICF) als een kader voor de behandeling van alle domeinen die verband houden met de gezondheid, (3) zorg voor gezamenlijke besluitvorming, en (4) gebruik patiëntgerapporteerde uitkomstmaten. Dit onderzoek heeft geleid tot inzicht in belemmerende en bevorderende factoren die ons kunnen helpen om strategieën te ontwikkelen om persoonsgerichte zorg te verbeteren. We hebben 14 aanbevelingen opgesteld voor de richtlijnontwikkelaars, personen die de richtlijn implementeren en gebruikers van de richtlijn.

**Hoofdstuk 6** rapporteert de bevindingen van een observationele cohortstudie over het gebruik van de patiëntgerapporteerde uitkomstmaten (PROMs) voor het stellen van doelen en het inzichtelijk maken van de uitkomst. Het routinematig gebruik van PROMs is van toenemend belang, omdat PROMs kunnen ondersteunen in de diagnostiek, het gezamenlijk stellen van belangrijkste behandeldoelen voor de patiënt en om de uitkomsten van de behandeling te evalueren. De doelen van deze studie waren 1) het inventariseren van het gebruik van PROMs bij patiënten met nekpijn en lage rugpijn; 2) het in kaart brengen in hoeverre de doelstellingen overeenkomen met de gekozen PROMs; 3) het verkennen wat de uitkomsten op basis van PROMs-metingen zijn.

In totaal hebben 60 fysiotherapeuten uit 43 praktijken deelgenomen. Alle informatie van de behandelepisode van elke patiënt werd geregistreerd in het ontwikkelde elektronische patiëntendossier. De geselecteerde PROMs waren de Visual Analogue Scale (VAS) voor pijnklachten en voor de beperkingen in activiteiten en participatie de Patiënt Specifieke Klachten (PSK), de Quebec Back Pain Disability Scale (QBPDS), en de Neck Disability Index (NDI). De fysiotherapeuten stelden de doelen op basis van de gegevens uit de anamnese (met inbegrip van de scores op de PROMs), onderzoek en de diagnose. Deze doelstellingen werden vastgelegd op een voorgestructureerd formulier op basis van een verkorte lijst van categorieën van de International Classification of Functioning, Disability and Health (ICF). We analyseerden de gegevens van 299 patiënten met nekpijn of lage rugpijn. Bij 70% van de patiënten met nekpijn was een enkele meting afgenomen, en bij 61% van de patiënten een herhaalde meting. Bij 68% van de patiënten met lage rugpijn was een enkele meting afgenomen en een herhaalde meting bij 63% van de patiënten. Bij patiënten met een doel gericht op pijnreductie werd in 46% een VAS gebruikt, terwijl bij 43% van de patiënten met een doel gericht op het verbeteren van een activiteit of participatie een PSK, QBPDS of een NDI was afgenomen. Het grote aantal behandeldoelen (mediaan 7 per patiënt) verklaart de overlap tussen de verschillende items, waardoor het moeilijk is de belangrijkste doelen te selecteren en te beoordelen welke PROM het meest geschikt. Het gemiddelde verschil tussen de voor- en nameting op de PROMs was statistisch significant en de meeste patiënten lieten een klinisch relevante verbetering zien. We concludeerden dat PROMs werden gebruikt bij de meeste patiënten, echter, de overeenkomst met de gestelde doelen was matig. Aanvullend onderzoek is nodig om te bepalen hoe PROMs worden gebruikt bij van het stellen van doelen en hoe de uitkomsten op de PROMs kunnen worden gebruikt.

**Hoofdstuk 7** bevat een algehele discussie over de studies in dit proefschrift. De resultaten worden in breder perspectief geplaatst en vergeleken met andere studies. Om de kwaliteit van de fysiotherapeutische zorg te verbeteren worden verschillende aanbevelingen gedaan voor fysiotherapeuten, beleidsmakers en richtlijnontwikkelaars. De meest relevante gevolgen voor richtlijnontwikkeling en -implementatie worden beschouwd. Ook worden gevolgen voor de fysiotherapeuten en patiënten beschreven en worden aanbevelingen voor toekomstig onderzoek geschetst.

Om de kwaliteit en veiligheid van de fysiotherapeutische zorg te verbeteren is het belangrijk dat fysiotherapeuten systematisch reflecteren op hun handelen. Op deze manier kunnen hiaten in de kennis, het klinisch redeneren en vaardigheden worden geïdentificeerd. Hiervoor dienen fysiotherapeuten een houding van kritische reflectie te ontwikkelen, en te zorgen voor een routine van zelfreflectie tijdens het dagelijkse werk. Richtlijnen kunnen ondersteunend zijn in het proces van klinisch redeneren. Fysiotherapeuten dienen een positieve houding aan te nemen ten aanzien van

permanente beroepsontwikkeling, scholing en het gebruik van richtlijnen. Om ervoor te zorgen dat de gegevens in patiëntendossiers leiden tot inzicht in het gedrag van de fysiotherapeut, is het van belang dat de beleidsmakers meer aandacht besteden aan de ondersteuning bij het verzamelen en delen van informatie tussen de patiënt en de fysiotherapeut en het leveren van feedback op de uitkomsten. Dit bevordert het inzicht in de effectiviteit van de behandeling en draagt bij aan veilige zorg. Richtlijnontwikkelaars dienen tools toe te voegen aan richtlijnen waarmee het klinisch redeneren verder wordt ondersteund. PROMs kunnen de interactie tussen de patiënt en de fysiotherapeut bevorderen voor het stellen van doelen voor de patiënt en het evalueren van de behandeling. Patiëntversies van richtlijnen kunnen de betrokkenheid van de patiënt verbeteren. Ten slotte is richtlijnimplementatie van cruciaal belang voor het bevorderen van evidence-based practice.

# Dankwoord

Dank zij de hulp en enthousiasme van velen is dit proefschrift tot een goed einde gebracht. Ik kijk terug op een leerzame, inspirerende en intensieve tijd waarvan het resultaat voor u ligt. Een promotietraject uitvoeren in combinatie met een druk gezinsleven met jonge kinderen was niet altijd een makkelijke opgave, maar mijn motivatie heeft geen moment ter discussie gestaan en de voldoening is groot.

Ik wil graag de gelegenheid nemen een aantal mensen expliciet te bedanken voor de totstandkoming van dit proefschrift. Echter, mijn dank gaat tevens uit naar ieder die ik niet specifiek noem, maar wel zijn of haar bijdrage heeft geleverd aan het onderzoek of mij heeft gesteund.

Om te beginnen mijn promotor en copromotoren prof. dr Ria Nijhuis, dr Philip van der Wees, dr. Bart Staal en dr Jozé Braspenning. Dit team omvatte een gezonde mix aan diversiteit, creativiteit en complementariteit, zowel in persoonlijkheden als in visies.

Ria, ik wil je bedanken voor je begeleiding en het in mij gestelde vertrouwen. Je hebt me kansen geboden en ik heb heel veel van je geleerd. Je stond altijd voor me klaar, gaf altijd grondige feedback en je hebt me geleerd een kritische wetenschappelijke houding aan te nemen. Je scherpe blik, creativiteit en zorgzaamheid zijn van grote betekenis geweest voor mij. Ik prijs me gelukkig dat ik nog met je blijf werken.

Philip, mijn dank is groot voor je begeleiding. Vooral in de laatste 2 jaar hebben we intensief contact gehad en de rust en het optimisme dat je uitstraalt zijn een mooi voorbeeld voor me geweest. Je kwam altijd met waardevolle feedback, iets wat de kwaliteit van mijn werk ten goede kwam. Ik heb veel van je geleerd en ik kijk met veel plezier terug aan de congressen die we samen hebben bezocht.

Bart, dankjewel voor je opbouwend kritische blik en de fijne samenwerking. Je deur stond altijd open voor me en je inhoudelijke en methodologische kennis was groot. Daar heb ik veel gebruik van kunnen maken. Je wist feilloos de zwakke plekken in het onderzoek te detecteren en kon me goed op weg helpen als ik even de draad kwijt was. Jouw relativeringsvermogen, kunde en rust zijn onmisbaar geweest.

Jozé, je bijdrage tijdens de discussies die we hadden over mijn onderzoeken was altijd waardevol. Als enige 'niet-fysiotherapeut' had je een nuttige bijdrage door hardop je af te vragen waarom 'in de wereld van de fysiotherapie bepaalde keuzes worden gemaakt'. Dit was heel verfrissend (en wij hadden er ook niet altijd antwoord op). Je kon structuur aanbrengen door scherpe vragen te stellen en je praktische benadering hebben mij goed geholpen. Veel dank voor je inzet.

Verder wil ik alle medewerkers van IQ healthcare en mijn collega's van de G-I-N Allied Health Working Group bedanken voor hun inzet bij de totstandkoming van de onderzoeksgegevens. Speciaal wil ik mijn overige co-auteurs van mijn artikelen bedanken

voor hun bijdrage, met name Margot Tacken voor haar onmisbare hulp bij het schrijven van mijn eerste artikel en Marjo Maas voor haar onuitputtelijke kennis van onderwijs(theorieën). Ik heb veel van jullie geleerd.

Een speciaal woord van dank voor iedereen die de afgelopen jaren meewerkte aan het beschikbaar stellen van de dossiers, vragenlijsten, registraties en de interviews. Zonder deze bereidheid hadden de onderzoeken niet uitgevoerd kunnen worden.

Rob Oostendorp, je hebt mede aan de basis gestaan van mijn eerste onderzoek, wat heeft geleid tot mijn promotietraject. Je bent een inspiratiebron voor me geweest en ik wil je danken voor je vertrouwen en steun.

Prof. dr. P. Assendelft, prof. dr. R. de Bie en prof. dr. C. Veenhof. Jullie vormden samen de manuscriptcommissie die dit proefschrift beoordeeld en goedgekeurd heeft. Mijn hartelijke dank hiervoor.

Mijn collega's Hilly Calsbeek, Anita Huis, Nicole Ketelaar en Maud Heinen. Als laatste van onze kamer ben ik nu dan ook gepromoveerd. Jullie waren hele fijne kamergenoten en we hebben een goede tijden gehad. Wat heb ik genoten van jullie gezelligheid, persoonlijke aandacht, chocola en lekkere koffie, en het plezier dat we hadden toen Hilly, Nicole en ik gelijktijdig zwanger waren. Ik hoop dat onze R&A etentjes nog lang doorgaan! Betsie van Gaal en Getty Huisman, wat fijn dat jullie zijn aangeschoven bij deze etentjes en dank voor jullie belangstelling en adviezen in de laatste fase.

Hilly, we zijn bijna gelijk begonnen bij IQ healthcare. We hebben lief en leed gedeeld en zijn altijd samen gebleven ondanks de vele verhuizingen die er zijn geweest. Ik heb veel gehad aan je luisterend oor, praktische tips en de gezellige gesprekken. Ik ben heel blij dat je mijn paranimf wil zijn.

Mijn A-team collega's van het MKIB project Marjo Maas, Guus Meerhoff, Juliette Cruijsberg, Janine Liefers, Femke Driehuis, Annick Bakker-Jacobs. Dank voor jullie support en voor het fijne gezelschap wanneer we weer eens samen midden in de nacht een uur lang voor een brug in Zeeland of in een afgelegen dorp in België stonden (en de navigatie de schuld gaven). Guus, dank voor de gezelligheid tijdens onze autoritten en bij congressen. Je hebt me geleerd hoe wetenschap en beleid bij elkaar komen en ik kijk uit naar jouw proefschrift.

Ik ben in de afgelopen jaren vaak verhuisd en heb vele fijne kamergenoten gehad. Ik wil hen allemaal bedanken voor de gezelligheid, inspiratie en wijze lessen in de afgelopen jaren.

Ook de collega's van het KNGF wil ik bedanken voor hun interesse en betrokkenheid, met name Jesper Knoop, Karin Heijblom en Annemarie Trompert .

Naast werk is er nog zoveel meer. Mijn vrienden, in het bijzonder mijn studiegenoten Chantal en Mariken en schoolvriendinnen Kitty, Simone, Marlies, Eveline, Adelin, Ingrid.

De borrels, etentjes, theater- en saunabezoeken zorgden voor de broodnodige afleiding. Marieke, Nicole en Judith, dank voor jullie interesse en gezellige etentjes. Lotty, dank voor je sportieve steun.

Mijn familie. Bedankt voor jullie onvoorwaardelijke steun, liefde en interesse. Het is fijn te weten dat jullie altijd achter me staan. Lieve pap en mam, ik ben hartstikke trots op jullie. Door de mentaliteit van hard werken, doorzetten en klaar staan voor een ander die ik met de paplepel kreeg ingegoten, heb ik dit proefschrift kunnen afronden.

Bianca, mijn paranimf en grote zus. Ik leer veel van je en je enthousiasme, gedrevenheid en lach zijn aanstekelijk. Ik weet dat ik op je kan rekenen, je bent een echte organisator en denkt overal aan, alles wat een goede paranimf nodig heeft. Mijn zus Natasja, dank voor het meelezen en dat ik af en toe bij je mocht komen om me even af te sluiten om te schrijven. Dit was efficiënt, maar ook heel gezellig.

Mieke en Stef, ik kan me geen betere schoonouders wensen. Dank voor jullie betrokkenheid en onophoudelijke ondersteuning bij de organisatie van mijn drukke leven en de vrolijkheid en ontspanning naast mijn promotie.

Mijn kinderen Connie, Felix en Thijmen. Connie, je nieuwsgierigheid en levenslust zijn geweldig en geven me veel energie en plezier. Ondanks dat hetgeen ik doe voor jou soms niet duidelijk was en je vond dat ik af en toe veel weg was zeg je vaak tegen mensen als ze je vragen wat je later wil worden “hetzelfde als mama, artikelen schrijven”. Mijn lieve ‘mannen’ Felix en Thijmen; jullie kwamen in mijn leven toen ik volop bezig was met mijn promotie. Ik geniet van jullie heerlijke koppen, energie en frisse geest. Jullie aanwezigheid heeft mij geleerd te relativeren en ik ben trots op jullie.

Liefste Ivo, mijn leven en liefde deel ik met jou. Het is een feest om samen met jou te mogen zijn. Jouw geduld, optimisme, relativeringsvermogen en gezelligheid zijn belangrijk voor me, niet alleen bij het schrijven van dit proefschrift, maar vooral in het dagelijks leven. Je zorgde ervoor dat ik niet vergat om af en toe te ontspannen en je was mijn steun en toeverlaat de afgelopen jaren. Ik zal nu een paar stapjes extra doen, dit deed jij de afgelopen jaren voor mij.

# Over de auteur

Simone van Dulmen (1975) startte in 1993 met de studie fysiotherapie aan de Hogeschool van Arnhem en Nijmegen. Zij studeerde in 1997 af en werkte vanaf 1997 tot 2007 als fysiotherapeut in een eerstelijns fysiotherapiepraktijk 'Fysiotherapie en manuele therapie Bouten' in Arnhem. In 2002 begon zij aan de studie Gezondheidswetenschappen aan de Universiteit van Maastricht en behaalde in 2007 haar doctoraal binnen de leergang Arbeid en Gezondheid.

In 2007 startte zij bij het Scientific Center for Quality of Healthare (IQ healthcare) van het Radboudumc in Nijmegen. Naast haar promotieonderzoek heeft ze bij IQ healthcare gewerkt aan verschillende advies- en evaluatieopdrachten. Zo was ze betrokken bij de uitvoering van het programma Zichtbare Zorg en is ze sinds 2013 verbonden aan het KNGF programma 'Kwaliteit in Beweging'. Vanaf 2015 werkt ze als projectleider bij het NFU programma 'Doen of laten. Terugdringen van onnodige zorg'.

Van 2014 tot 2015 heeft ze daarnaast als beleidsmedewerker bij de afdeling Beleid en Ontwikkeling van het KNGF gewerkt. Ze is sinds 2015 voorzitter van de Allied Health Community van het Guidelines International Network (G-I-N).

Simone woont samen met Ivo Hendriks. Zij hebben samen drie kinderen, Connie (2006), Felix (2008) en Thijmen (2010).

# PhD portfolio

Institute for Health Sciences
Radboudumc

**Name PhD student:** S. A. van Dulmen
**Department:** IQ healthcare
**Graduate School:** Radboud Institute for Health Sciences

***PhD period:***
01-01-2009 – 11-02-2016
**Promotor(s):**
Prof. dr. MWG Nijhuis-van der Sanden
**Co-promotor(s):**
dr. PJ van der Wees
dr. JB Staal
dr. JCC Braspenning

| | Year(s) | ECTS |
|---|---|---|
| ***TRAINING ACTIVITIES*** | | |
| a) *Courses & Workshops* | | |
| - Leiding geven aan projecten, ICM | 2014 | 0.5 |
| - BROK (Certificate for Good Clinical Practice) | 2013 | 1.5 |
| - Multilevel analyse, EMGO | 2011 | 1.5 |
| - Schrijven van subsidieaanvragen, Zonmw | 2011 | 0.1 |
| - Implementation course, G-I-N | 2013 | 1.5 |
| - Qualitative research methods in healthcare (CaRe) | 2012 | 1.0 |
| - Presenteren eigen onderzoek, Radboud University | 2010 | 1.5 |
| - Schrijven van wetenschappelijke teksten, Radboud in'to languages | 2010 | 2.0 |
| - Academic writing, Radboud in'to languages | 2009 | 3.0 |
| - Cursus SPSS, PAO Heyendaal | 2009 | 0.5 |
| b) *Seminars & lectures* | | |
| - Seminar RGF, oral presentation | 2016 | 0.1 |
| - Seminar HAN, oral presentation | 2013 | 0.1 |
| - Seminar KNGF, oral presentation | 2010 | 0.1 |
| c) Symposia & congresses | | |
| - World Congress Physical Therapy (WCPT) (Singapore), 2 poster presentations and 1 oral presentation | 2015 | 2.0 |
| - Guidelines International Network (G-I-N) congres (Amsterdam), 2 oral presentations | 2015 | 2.0 |
| - Guidelines International Network (G-I-N) congres (Berlijn), 2 poster presentations | 2014 | 1.5 |
| - NFU symposium, oral presentation | 2014 | 0.2 |
| - Jaarcongres KNGF, 4 oral presentations | 2011-2015 | 2.0 |
| - NCEBP (RIHS) Symposium, oral presentation | 2010, 2012 | 0.2 |
| - ISQUA Conference (Zwitserland), oral presentation | 2012 | 1.5 |
| - Low Back Pain Forum Odense (Denemarken), oral presentation | 2012 | 0.5 |
| - Guidelines International Network (G-I-N) congres (Korea), workshop | 2010 | 2.0 |
| - World Congress Physical Therapy (WCPT) (Amsterdam), oral presentation | 2010 | 1.0 |

| | | |
|---|---|---|
| d) *Other* | | |
| - Monthly researchers meeting, IQ healthcare | 2010-2015 | 2.0 |
| - Member expert group transparency, IQ healthcare | 2010-2015 | 2.0 |
| - Review scientific publications, several journals | 2013-2016 | 0.5 |

***TEACHING ACTIVITIES***

| | | |
|---|---|---|
| e) *Lecturing* | | |
| - Lecturing Bachelor Biomedical sciences Radboud university | 2013-2014 | 0.1 |
| f) *Supervision of internships / other* | | |
| - Supervision internship bachelor students Radboud university | 2013-2016 | 1.0 |

# List of publications

## Scientific Publicaties (* listed in this thesis)

2016

*__van Dulmen SA__, van der Wees PJ, Staal JB, Braspenning JCC, Nijhuis-van der Sanden MWG. Patient reported outcome measures (PROMs) for goalsetting and outcome measurement in primary care physiotherapy, an explorative field study. Physiotherapy (2016, in press).

2015

*__van Dulmen SA__, Lukersmith S, Muxlow J, Santa Mina E, Nijhuis-van der Sanden MW, van der Wees PJ on behalf of the G-I-N Allied Health Steering Group. Supporting a person-centred approach in clinical guidelines. A position paper of the Allied Health Community – Guidelines International Network (G-I-N). Health Expectations 2015; 18(5): 1543-1558.

*Maas MJ, __van Dulmen SA__, Sagasser MH, Heerkens YF, van der Vleuten CP, Nijhuis-van der Sanden MW, van der Wees PJ. Critical features of peer assessment of clinical performance to enhance adherence to a low back guideline for physical therapists: a mixed methods design. BMC Medical Education 2015; 15(1): 203

2014

*__van Dulmen SA__, Maas M, Staal JB, Rutten G, Kiers H, Nijhuis-van der Sanden M, van der Wees P Effectiveness of peer assessment for implementing a Dutch physical therapy low back pain guideline: cluster randomized controlled trial. Phys Ther. 2014 Oct;94(10):1396-409.

__Simone van Dulmen__, Marjo Maas, Bart Staal, Geert Rutten, Henri Kiers, Maria Nijhuis-van der Sanden, Philip van der Wees. Effectiveness of Peer Assessment for Implementing a Dutch Physical Therapy Low Back Pain Guideline: A Cluster Randomized Controlled Trial. Fysiopraxis 2014, 11 (23).

2013

Martijn L, Harmsen M, Gaal S, Mettes D, __van Dulmen S__, Wensing M. Are health professionals' perceptions of patient safety related to figures on safety incidents? J Eval Clin Pract. 2013 Oct;19(5):944-947.

2011

*__van Dulmen, Simone A__.; Tacken, Margot A.J.B.; Staal, J. Bart; Gaal, Sander; Wensing, Michel; Nijhuis-van der Sanden, Maria W.G. Patient Safety in Primary Allied Health Care: What Can We Learn From Incidents in a Dutch Exploratory Cohort Study? Medical Care. 2011, 49(12):1089-1096.

**Dulmen van S**, Tacken M, Staal JB, et al. Incidenten in eerstelijns paramedische praktijken. Ned Tijdschr Fysiother 2011, 121; 162-175

2010

M. Harmsen, S. Gaal, **S. van Dulmen**, E. de Feijter, P. Giesen, J. Jacobs, L. Martijn, D. Mettes, W. Verstappen, R. Nijhuis-van de Sanden, M.J.P. Wensing, Patient safety in Dutch primary care: a study protocol. Implementation Science 2010; 5:50.

M. Tacken, **S. van Dulmen**, M.J.P. Wensing, R. Nijhuis-van de Sanden. Patientveiligheid in eerstelijns oefentherapie praktijken. Beweegreden 2010,6, (2); 22-23.

2008

R. Nijhuis-van de Sanden, G. Rutten, **S. van Dulmen**, C.W.M. Neeleman-Van der Steen. Het Fysiotherapeutisch Elektronisch Patiëntendossier: kans of bedreiging voor de klinimetrie? RGF Bulletin 2008.

## Reports

2016

G. Meerhoff, **S. van Dulmen**, J. Cruijsberg, J. Liefers, P van der Wees, R. van der Sanden. Eindrapport uitkomsten van zorg binnen het Rug-netwerk. Pilot binnen het onderzoeksprogramma Masterplan Kwaliteit in Beweging. Eindrapport. Nijmegen, IQ healthcare, 2016.

2015

**S. van Dulmen**, P van der Wees, G. Meerhoff, M. Maas, R. van der Sanden. Eindrapport uitkomsten van zorg binnen de coöperatie Fysiocare. Pilot binnen het onderzoeksprogramma Kwaliteit in beweging. Eindrapport. Nijmegen, IQ healthcare, 2015.

**S. van Dulmen**, G. Meerhoff,M. Maas, P van der Wees, R. van der Sanden. Eindrapport uitkomsten van zorg binnen de coöperatie Fysio-Omni. Pilot binnen het onderzoeksprogramma Kwaliteit in beweging. Eindrapport. Nijmegen, IQ healthcare, 2015.

**S. van Dulmen**, M. Maas, P van der Wees, R. van der Sanden. Eindrapport uitkomsten van zorg binnen de coöperatie Fysio10IJsselland Pilot binnen het onderzoeksprogramma Kwaliteit in beweging. Eindrapport. Nijmegen, IQ healthcare, 2015.

2013

M. Wensing, **S. van Dulmen**. Retirement of quality indicators. A quick scan for the German SQG program. Rapport. Nijmegen, IQ healthcare, 2013.

**S. van Dulmen**. Ondersteuning programma Achmea Kwaliteit van Zorg starterspakketten: lage rug. Rapport. Nijmegen, IQ healthcare, 2013.

**S. van Dulmen**, A. Oerlemans. Ondersteuning programma Achmea Kwaliteit van Zorg starterspakketten: Parkinson. Rapport. Nijmegen, IQ healthcare, 2013.

**S. van Dulmen**. Ondersteuning programma Achmea Kwaliteit van Zorg starterspakketten: Verstandelijke beperking. Rapport. Nijmegen, IQ healthcare, 2013.

2012

**S. van Dulmen**, H. Calsbeek, J. Koetsenruijter, J. Braspenning. Kwaliteitsindicatoren Eerstelijns Fysiotherapie (Kwaliefy); Uitvraag en gebruik 2011. Rapport. Nijmegen, IQ healthcare, 2012.

W. Verstappen, K. Chun Cheung, **S. van Dulmen**, P. Giesen, P. van Grunsven, A. van Holten-Verzantvoort, D. Mettes, R. Nijhuis-van der Sanden, D. Opstelten, M. Wensing. Tweede Nationale Meldweek Patientveiligheid eerstelijn. Rapport. Nijmegen, IQ healthcare, 2012.

2009

M. Tacken, **S. van Dulmen**, R. Nijhuis-van de Sanden. Patiëntveiligheid in de Nederlandse eerstelijnszorg anno 2009: resultatendeelonderzoek paramedische praktijken. Rapport. Nijmegen, IQ healthcare, 2009.

C.W.M. Neeleman-Van der Steen, J. Braspenning, E. Hendriks, M. Nijkrake, G. Rutten, **S. van Dulmen**, R. Nijhuis-van de Sanden, R. de Bie, R.A.B. Oostendorp. Ontwikkeling kwaliteitsindicatoren voor KNGF richtlijnen. Een handleiding. Rapport. Nijmegen, IQ healthcare, 2009.

**S. van Dulmen**, R.A.B. Oostendorp, R. Nijhuis-van de Sanden, Project Kwaliteitsindicatoren Fysiotherapie 3 (ProKwaF-3). Kwaliteit Fysiotherapeutische Verslaglegging. Rapport. Nijmegen, IQ healthcare, 2009.

Institute for Health Sciences
Radboudumc